APPLICATION FOR

UNITED STATES LETTERS PATENT

This application claims benefit of co-pending U.S. Patent Application Serial No. 60/264,649 filed January 26, 2001, entitled "Mosquito Olfactory Genes, Polypeptides, and Methods of Use Thereof" which is hereby incorporated by reference. Be it known that I, Laurence J Zwiebel, a citizen of the United States, residing at 2512 Sunset Place, Nashville, TN 37212; have invented a new and useful "Mosquito Olfactory Genes, Polypeptides, and Methods of Use Thereof".

GOVERNMENT SUPPORT CLAUSE

This invention was made with federal grant money under NIH grant 1 R01 DC04692-01 and NSF grant 0075338. The United States Government has certain rights in this invention.

in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

FIELD OF THE INVENTION

The present invention relates generally to the field of host identification by insects. Specifically, the present invention relates to the identification and cloning of genes related to mosquito olfaction, identification and purification of polypeptides thereof, and methods of use thereof.

BACKGROUND OF THE INVENTION

The ability of an insect to respond to chemical stimuli is necessary for the insect to reproduce, mate, and feed. For example, insects respond to certain chemical stimuli by moving up a chemical gradient to identify and target a host. Mosquitoes, in particular, are believed to use olfaction to identify and target sources of bloodmeal for reproductive purposes. This behavior contributes to the spread of diseases in humans, such as malaria, encephalitis, and dengue fever; as well as, animal and livestock disease.

Olfaction plays a critical role in insect behaviors among agricultural pests and disease vectors. Hildebrand, et al., 1997, Annu. Rev. Neurosci, 20:595-631. In *Drosophila melanogaster* (the common fruit fly), the olfactory system functions through a rapid cycling between an on and off state of certain regulatory molecules.

The olfactory signal transduction cascade is "turned on" by ligand-based activation of an odorant receptor and transduction of the signal by G-protein coupled second messenger pathways Boekhoff *et al.*, 1994, J. Neurosci, 14:3304-9. The "on signal" is rapidly and substantially terminated in the *Drosophila* system through the modification of the odorant receptor such that the G-protein coupled second messenger pathway is deactivated. Dohlman *et al.*, 1991, Annual Review of Biochemistry, 60:653-88. Olfactory transduction is provided by second messenger pathways of G protein-coupled receptors. Reed, R., 1992, Neuron 8:205-209; Bloekhoff, *et al*, 1994, Neurosci 14:3304-3309.

The structural and functional characteristics of the mosquito olfactory system has not been characterized to date. Given the importance of the controlling this pest and disease vector, what is needed is the identification and characterization of the genes and polypeptides that function for mosquito olfaction and methods of use thereof for mosquito management.

SUMMARY OF THE INVENTION

The present invention provides, in part, eight novel mosquito polypeptides and nucleic acids encoding the polypeptides (collectively referred to herein as "mosquito olfaction molecules"). Seven of the polypeptides are novel mosquito odorant receptors and the eighth is a novel mosquito arrestin molecule (see Figure 8). The odorant receptor molecules are discovered to function in a ligand-induced signal transduction

pathway for the activation of mosquito olfaction. The mosquito arrestin molecule is discovered to function to inhibit the activated signal transduction cascade. Thus, the odorant receptors can be viewed as parts of an "on switch" or an "on signal" and the arrestin molecule can be viewed as an "off switch" or an "off signal" for the odorant detection system of the mosquito. The present invention is not bound by theory or mechanism.

The present invention also provides, in part, a system for disrupting the mosquito olfactory system by disrupting, inhibiting, or otherwise interfering with the function of the off switch for mosquito olfaction. Such interference is contemplated to inhibit or degrade the ability of the mosquito to appropriately respond to chemical clues in the environment used by the mosquito for host identification and targeting. For, example, if the signal cascade cannot be terminated or inhibited, then the mosquito is impaired in following a chemical gradient to a host through sampling of the frequency of ligand-induced activation of the olfaction signal cascade. In this example, the chemical concentration of the odorant is expected to increase with decreasing distance to the target. Thus, receptor activation is expected to increase with decreasing distance to the target. It is a discovery of the present invention, that factors that inhibit the on and off cycling of the mosquito olfactory signal cascade through inhibition of signal deactivation are useful for the control of mosquitoes. Test agents used in a method for identifying mosquito olfaction molecule binding compounds would include, but

are not limited to: chemicals, proteins, peptides, organic compounds and lipids. Such factors that inhibit signal deactivation may be peptides and chemicals. Several classes of chemicals that would be selected as targets are the carboxylic acids and steroids that are components of human sweat. Cork, A. (1996). Olfactory sensing is the basis of host location by mosquitoes and other hematophagous Diptera. In Olfaction in Mosquito-Host Interactions, G. R. B. a. G. Cardew, ed. (Chichester, New York, Brisbane, Toronto , Singapor: John Wiley & Sons), pp. 71-84. Furthermore, certain aspects of the present invention are contemplated to be effective for insects in general.

Methods are presented for identifying compounds that interfere with the operation of the mosquito olfactory system resulting in an over stimulation of olfactory signaling. One consequence of interfering with the mosquito olfactory system is that the mosquito has a diminished ability to home in on sources of bloodmeal. Additionally, interfering with mosquito insect olfactory systems will inhibit mating and feeding having a significant impact on mosquito populations and is helpful, for example, in nuisance and disease vector control for humans and livestock. Interfering with non-mosquito insect olfaction will similarly have a positive impact in control of other insect populations including for the protection of crops, such as: wheat, corn, rice, cotton, and soybeans. Thus, certain aspects of the present invention provide screening assays for the identification of compositions that will reduce the ability of mosquitoes to locate sources of bloodmeal, such as humans and other mammals,

including livestock (cattle, pigs, horses, sheep, etc.), show animals (horses, pigs, sheep, dogs, cats, etc.), and pets (dogs, cats, horses, etc). Certain aspects of the present invention provide a screening assay for the production of "mosquito olfaction molecules."

One aspect of the present invention provides an isolated DNA comprising a nucleotide sequence that encodes arrestin 1 polypeptide (*e.g.*, SEQ ID NO: 2). In certain embodiments, arrestin 1 nucleotide sequence comprises a DNA molecule that hybridizes under stringent conditions to a DNA having a nucleotide sequence consisting of SEQ ID NO: 1, or the complement of SEQ ID NO: 1. Preferably the isolated DNA encodes naturally-occurring *Anopheles gambiae* arrestin 1 polypeptides. In certain embodiments, the nucleotide sequence may be that of SEQ ID NO: 1. In alternate embodiments, the nucleotide sequence may encode a fragment of SEQ ID NO: 2 at least 20 residues in length. One of ordinary skill in the art knows that a polypeptide fragment having a length of 20 residues is capable of functioning as an immunogen. In certain embodiments, the nucleotide sequence may encode a polypeptide having a conservatively modified amino acid sequence of SEQ ID NO: 2. In certain embodiments, the isolated polynucleotide comprises a complement to a sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 2, and conservatively modified SEQ ID NO: 2. In alternate embodiments, the nucleotide sequence may be that of degenerate variants of above-mentioned sequences. The invention also includes operably linking one or

more expression control sequences to any of the above-mentioned nucleotide sequences. The invention also includes a cell comprising any of the above-mentioned nucleotide sequences operably linked to one or more expression control sequences.

The present invention also provides an isolated DNA comprising a nucleotide sequence that encodes odorant receptor 1 polypeptide (*e.g.*, SEQ ID NO: 4). In certain embodiments, odorant receptor 1 nucleotide sequence comprises a DNA molecule that hybridizes under stringent conditions to a DNA having a nucleotide sequence consisting of SEQ ID NO: 3, or the complement of SEQ ID NO: 3. Preferably the isolated DNA encodes naturally-occurring *Anopheles gambiae* odorant receptor 1 polypeptides. In certain embodiments, the nucleotide sequence may be that of SEQ ID NO: 3. In alternate embodiments, the nucleotide sequence may encode a fragment of SEQ ID NO: 4 at least 20 residues in length. One of ordinary skill in the art knows that a polypeptide fragment having a length of 20 residues is capable of functioning as an immunogen. In certain embodiments, the nucleotide sequence may encode a polypeptide having a conservatively modified amino acid sequence of SEQ ID NO: 4. In certain embodiments, the isolated polynucleotide comprises a complement to a sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 4, and conservatively modified SEQ ID NO: 4. In other alternate embodiments, the nucleotide sequence may be that of degenerate variants of above-mentioned sequences. The invention also includes operably linking one or more expression control sequences to any of the above-mentioned nucleotide

sequences. The invention also includes a cell comprising any of the above-mentioned nucleotide sequences operably linked to one or more expression control sequences.

The present invention provides an isolated DNA comprising a nucleotide sequence that encodes odorant receptor 2 polypeptide (*e.g.*, SEQ ID NO: 6). In certain embodiments, odorant receptor 2 nucleotide sequence comprises a DNA molecule that hybridizes under stringent conditions to a DNA having a nucleotide sequence consisting of SEQ ID NO: 5, or the complement of SEQ ID NO: 5. Preferably the isolated DNA encodes naturally-occurring *Anopheles gambiae* odorant receptor 2 polypeptides. In certain embodiments, the nucleotide sequence may be that of SEQ ID NO: 5. In alternate embodiments, the nucleotide sequence may encode a fragment of SEQ ID NO: 6 at least 20 residues in length. One of ordinary skill in the art knows that a polypeptide fragment having a length of 20 residues is capable of functioning as an immunogen. In certain embodiments, the nucleotide sequence may encode a polypeptide having a conservatively modified amino acid sequence of SEQ ID NO: 6. In certain embodiments, the isolated polynucleotide comprises a complement to a sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 6, and conservatively modified SEQ ID NO: 6. In other alternate embodiments, the nucleotide sequence may be that of degenerate variants of above-mentioned sequences. The invention also includes operably linking one or more expression control sequences to any of the above-mentioned nucleotide

sequences. The invention also includes a cell comprising any of the above-mentioned nucleotide sequences operably linked to one or more expression control sequences.

The present invention also provides an isolated DNA comprising a nucleotide sequence that encodes odorant receptor 3 polypeptide (*e.g.*, SEQ ID NO: 8). In certain embodiments, odorant receptor 3 nucleotide sequence comprises a DNA molecule that hybridizes under stringent conditions to a DNA having a nucleotide sequence consisting of SEQ ID NO: 7, or the complement of SEQ ID NO: 7. Preferably the isolated DNA encodes naturally-occurring *Anopheles gambiae* odorant receptor 3 polypeptides. In certain embodiments, the nucleotide sequence may be that of SEQ ID NO: 7. In alternate embodiments, the nucleotide sequence may encode a fragment of SEQ ID NO: 8 at least 20 residues in length. One of ordinary skill in the art knows that a polypeptide fragment having a length of 20 residues is capable of functioning as an immunogen. In certain embodiments, the nucleotide sequence may encode a polypeptide having a conservatively modified amino acid sequence of SEQ ID NO: 8. In certain embodiments, the isolated polynucleotide comprises a complement to a sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 8, and conservatively modified SEQ ID NO: 8. In other alternate embodiments, the nucleotide sequence may be that of degenerate variants of above-mentioned sequences. The invention also includes operably linking one or more expression control sequences to any of the above-mentioned nucleotide

sequences. The invention also includes a cell comprising any of the above-mentioned nucleotide sequences operably linked to one or more expression control sequences.

The present invention also provides an isolated DNA comprising a nucleotide sequence that encodes odorant receptor 4 polypeptide (*e.g.*, SEQ ID NO: 14). In certain embodiments, odorant receptor 4 nucleotide sequence comprises a DNA molecule that hybridizes under stringent conditions to a DNA having a nucleotide sequence consisting of SEQ ID NO: 13, or the complement of SEQ ID NO: 13. Preferably the isolated DNA encodes naturally-occurring *Anopheles gambiae* odorant receptor 4 polypeptides. In certain embodiments, the nucleotide sequence may be that of SEQ ID NO: 13. In alternate embodiments, the nucleotide sequence may encode a fragment of SEQ ID NO: 14 at least 20 residues in length. One of ordinary skill in the art knows that a polypeptide fragment having a length of 20 residues is capable of functioning as an immunogen. In certain embodiments, the nucleotide sequence may encode a polypeptide having a conservatively modified amino acid sequence of SEQ ID NO: 14. In certain embodiments, the isolated polynucleotide comprises a complement to a sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 14, and conservatively modified SEQ ID NO: 14. In other alternate embodiments, the nucleotide sequence may be that of degenerate variants of above-mentioned sequences. The invention also includes operably linking one or more expression control sequences to any of the above-mentioned nucleotide sequences. The invention also includes a cell comprising any of

the above-mentioned nucleotide sequences operably linked to one or more expression control sequences.

The present invention also provides an isolated DNA comprising a nucleotide sequence that encodes odorant receptor 5 polypeptide (*e.g.*, SEQ ID NO: 16). In certain embodiments, odorant receptor 5 nucleotide sequence comprises a DNA molecule that hybridizes under stringent conditions to a DNA having a nucleotide sequence consisting of SEQ ID NO: 15, or the complement of SEQ ID NO: 15. Preferably the isolated DNA encodes naturally-occurring *Anopheles gambiae* odorant receptor 5 polypeptides. In certain embodiments, the nucleotide sequence may be that of SEQ ID NO: 15. In alternate embodiments, the nucleotide sequence may encode a fragment of SEQ ID NO: 16 at least 20 residues in length. One of ordinary skill in the art knows that a polypeptide fragment having a length of 20 residues is capable of functioning as an immunogen. In certain embodiments, the nucleotide sequence may encode a polypeptide having a conservatively modified amino acid sequence of SEQ ID NO: 16. In certain embodiments, the isolated polynucleotide comprises a complement to a sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 16, and conservatively modified SEQ ID NO: 16. In other alternate embodiments, the nucleotide sequence may be that of degenerate variants of above-mentioned sequences. The invention also includes operably linking one or more expression control sequences to any of the above-mentioned nucleotide sequences. The invention also includes a cell comprising any of

the above-mentioned nucleotide sequences operably linked to one or more expression control sequences.

The present invention also provides an isolated DNA comprising a nucleotide sequence that encodes odorant receptor 6 polypeptide (*e.g.*, SEQ ID NO: 18). In certain embodiments, odorant receptor 6 nucleotide sequence comprises a DNA molecule that hybridizes under stringent conditions to a DNA having a nucleotide sequence consisting of SEQ ID NO: 17, or the complement of SEQ ID NO: 17. Preferably the isolated DNA encodes naturally-occurring *Anopheles gambiae* odorant receptor 6 polypeptides. In certain embodiments, the nucleotide sequence may be that of SEQ ID NO: 17. In alternate embodiments, the nucleotide sequence may encode a fragment of SEQ ID NO: 18 at least 20 residues in length. One of ordinary skill in the art knows that a polypeptide fragment having a length of 20 residues is capable of functioning as an immunogen. In certain embodiments, the nucleotide sequence may encode a polypeptide having a conservatively modified amino acid sequence of SEQ ID NO: 18. In certain embodiments, the isolated polynucleotide comprises a complement to a sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 18, and conservatively modified SEQ ID NO: 18. In other alternate embodiments, the nucleotide sequence may be that of degenerate variants of above-mentioned sequences. The invention also includes operably linking one or more expression control sequences to any of the above-mentioned nucleotide sequences. The invention also includes a cell comprising any of

the above-mentioned nucleotide sequences operably linked to one or more expression control sequences.

The present invention also provides an isolated DNA comprising a nucleotide sequence that encodes odorant receptor 7 polypeptide (*e.g.*, SEQ ID NO: 20). In certain embodiments, odorant receptor 7 nucleotide sequence comprises a DNA molecule that hybridizes under stringent conditions to a DNA having a nucleotide sequence consisting of SEQ ID NO: 19, or the complement of SEQ ID NO: 19. Preferably the isolated DNA encodes naturally-occurring *Anopheles gambiae* odorant receptor 7 polypeptides. In certain embodiments, the nucleotide sequence may be that of SEQ ID NO: 19. In alternate embodiments, the nucleotide sequence may encode a fragment of SEQ ID NO: 20 at least 20 residues in length. One of ordinary skill in the art knows that a polypeptide fragment having a length of 20 residues is capable of functioning as an immunogen. In certain embodiments, the nucleotide sequence may encode a polypeptide having a conservatively modified amino acid sequence of SEQ ID NO: 20. In certain embodiments, the isolated polynucleotide comprises a complement to a sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 20, and conservatively modified SEQ ID NO: 20. In other alternate embodiments, the nucleotide sequence may be that of degenerate variants of above-mentioned sequences. The invention also includes operably linking one or more expression control sequences to any of the above-mentioned nucleotide sequences. The invention also includes a cell comprising any of

the above-mentioned nucleotide sequences operably linked to one or more expression control sequences.

The present invention provides a substantially pure arrestin 1 polypeptide that includes amino acid sequence that contains at least a conservatively modified identity with SEQ ID NO: 2 and binds to odorant receptors. The amino acid sequence of arrestin 1 protein can differ from SEQ ID NO: 2 by non-conservative substitutions, deletions, or insertions located at positions that do not destroy the function of the arrestin 1 polypeptide. In alternate embodiments, the polypeptide has an amino acid sequence consisting of SEQ ID NO: 2. The purified polypeptide is a polypeptide that binds specifically to an antibody that binds specifically to mosquito arrestin. In other alternate embodiments, the polypeptide comprises fragments of SEQ ID NO: 2, having at least 20 consecutive residues.

The present invention also provides a substantially pure odorant receptor 1 polypeptide that includes amino acid sequence that contains at least a conservatively modified identity with SEQ ID NO: 4 and binds to arrestin. The amino acid sequence of odorant receptor 1 polypeptide can differ from SEQ ID NO: 4 by non-conservative substitutions, deletions, or insertions located at positions that do not destroy the function of the odorant receptor 1 polypeptide. In alternate embodiments, the polypeptide has an amino acid sequence consisting of SEQ ID NO: 4. In other alternate embodiments, the polypeptide comprises fragments of SEQ ID NO: 4, having at least 20 consecutive residues.

The present invention provides a substantially pure odorant receptor 2 polypeptide that includes amino acid sequence that contains at least a conservatively modified identity with SEQ ID NO: 6 and binds to arrestin. The amino acid sequence of odorant receptor 2 polypeptide can differ from SEQ ID NO: 6 by non-conservative substitutions, deletions, or insertions located at positions that do not destroy the function of the odorant receptor 2 polypeptide. In alternate embodiments, the polypeptide has an amino acid sequence consisting of SEQ ID NO: 6. In other alternate embodiments, the polypeptide comprises fragments of SEQ ID NO: 6, having at least 20 consecutive residues.

The present invention also provides a substantially pure odorant receptor 3 polypeptide that includes amino acid sequence that contains at least a conservatively modified identity with SEQ ID NO: 8 and binds to arrestin. The amino acid sequence of odorant receptor 3 polypeptide can differ from SEQ ID NO: 8 by non-conservative substitutions, deletions, or insertions located at positions that do not destroy the function of the odorant receptor 3 polypeptide. In alternate embodiments, the polypeptide has an amino acid sequence consisting of SEQ ID NO: 8. In other alternate embodiments, the polypeptide comprises fragments of SEQ ID NO: 8, having at least 20 consecutive residues.

The present invention also provides a substantially pure odorant receptor 4 polypeptide that includes amino acid sequence that contains at least a conservatively modified identity with SEQ ID NO: 14 and binds to arrestin. The amino acid sequence

of odorant receptor 4 polypeptide can differ from SEQ ID NO: 14 by non-conservative substitutions, deletions, or insertions located at positions that do not destroy the function of the odorant receptor 4 polypeptide. In alternate embodiments, the polypeptide has an amino acid sequence consisting of SEQ ID NO: 14. In other alternate embodiments, the polypeptide comprises fragments of SEQ ID NO: 14, having at least 20 consecutive residues.

The present invention also provides a substantially pure odorant receptor 5 polypeptide that includes amino acid sequence that contains at least a conservatively modified identity with SEQ ID NO: 16 and binds to arrestin. The amino acid sequence of odorant receptor 5 polypeptide can differ from SEQ ID NO: 16 by non-conservative substitutions, deletions, or insertions located at positions that do not destroy the function of the odorant receptor 5 polypeptide. In alternate embodiments, the polypeptide has an amino acid sequence consisting of SEQ ID NO: 16. In other alternate embodiments, the polypeptide comprises fragments of SEQ ID NO: 16, having at least 20 consecutive residues.

The present invention also provides a substantially pure odorant receptor 6 polypeptide that includes amino acid sequence that contains at least a conservatively modified identity with SEQ ID NO: 18 and binds to arrestin. The amino acid sequence of odorant receptor 6 polypeptide can differ from SEQ ID NO: 18 by non-conservative substitutions, deletions, or insertions located at positions that do not destroy the function of the odorant receptor 6 polypeptide. In alternate embodiments, the

polypeptide has an amino acid sequence consisting of SEQ ID NO: 18. In other alternate embodiments, the polypeptide comprises fragments of SEQ ID NO: 18, having at least 20 consecutive residues.

The present invention also provides a substantially pure odorant receptor 7 polypeptide that includes amino acid sequence that contains at least a conservatively modified identity with SEQ ID NO: 20 and binds to arrestin. The amino acid sequence of odorant receptor 7 polypeptide can differ from SEQ ID NO: 20 by non-conservative substitutions, deletions, or insertions located at positions that do not destroy the function of the odorant receptor 7 polypeptide. In alternate embodiments, the polypeptide has an amino acid sequence consisting of SEQ ID NO: 20. In other alternate embodiments, the polypeptide comprises fragments of SEQ ID NO: 20, having at least 20 consecutive residues.

The invention also provides an arrestin 1 antibody, which comprises polyclonal or monoclonal antibodies. The antibody can be conjugated to a detectable label.

Another aspect of the present invention provides an odorant receptor 1 antibody, which comprises polyclonal or monoclonal antibodies. The antibody can be conjugated to a detectable label. Antibody labels and methods are well known in the art.

The present invention also provides an odorant receptor 2 antibody, which comprises polyclonal or monoclonal antibodies. The antibody can be conjugated to a detectable label.

Another aspect of the present invention provides an odorant receptor 3 antibody, which comprises polyclonal or monoclonal antibodies. The antibody can be conjugated to a detectable label.

Another aspect of the present invention provides an odorant receptor 4 antibody, which comprises polyclonal or monoclonal antibodies. The antibody can be conjugated to a detectable label.

Another aspect of the present invention provides an odorant receptor 5 antibody, which comprises polyclonal or monoclonal antibodies. The antibody can be conjugated to a detectable label.

Another aspect of the present invention provides an odorant receptor 6 antibody, which comprises polyclonal or monoclonal antibodies. The antibody can be conjugated to a detectable label.

Another aspect of the present invention provides an odorant receptor 7 antibody, which comprises polyclonal or monoclonal antibodies. The antibody can be conjugated to a detectable label.

The present invention also presents a method of producing arrestin 1 protein. The method includes the following steps: (a) providing a cell transformed with an isolated DNA comprising a nucleotide sequence that encodes an amino acid sequence of SEQ ID NO: 2; (b) culturing the cell; and (c) collecting from the cell or the medium of the cell the polypeptide encoded by the polynucleotide sequence. Certain

alternatives to SEQ ID NO: 2 are described above (*e.g.* conservative variants and hybridization variants).

The present invention also provides a method of manufacturing odorant receptor 1 protein. The method includes the following steps: (a) providing a cell transformed with an isolated DNA comprising a nucleotide sequence that encodes an amino acid sequence of SEQ ID NO: 4; (b) culturing the cell; and (c) collecting from the cell or the medium of the cell the polypeptide encoded by the polynucleotide sequence.

The present invention provides a method of manufacturing odorant receptor 2 protein. The method includes the following steps: (a) providing a cell transformed with an isolated DNA comprising a nucleotide sequence that encodes an amino acid sequence of SEQ ID NO: 6; (b) culturing the cell; and (c) collecting from the cell or the medium of the cell the polypeptide encoded by the polynucleotide sequence.

The present invention also provides a method of manufacturing odorant receptor 3 protein. The method includes the following steps: (a) providing a cell transformed with an isolated DNA comprising a nucleotide sequence that encodes an amino acid sequence of SEQ ID NO: 8; (b) culturing the cell; and (c) collecting from the cell or the medium of the cell the polypeptide encoded by the polynucleotide sequence.

The present invention also provides a method of manufacturing odorant receptor 4 protein. The method includes the following steps: (a) providing a cell transformed with an isolated DNA comprising a nucleotide sequence that encodes an amino acid sequence of SEQ ID NO: 14; (b) culturing the cell; and (c) collecting from

the cell or the medium of the cell the polypeptide encoded by the polynucleotide sequence.

The present invention also provides a method of manufacturing odorant receptor 5 protein. The method includes the following steps: (a) providing a cell transformed with an isolated DNA comprising a nucleotide sequence that encodes an amino acid sequence of SEQ ID NO: 16; (b) culturing the cell; and (c) collecting from the cell or the medium of the cell the polypeptide encoded by the polynucleotide sequence.

The present invention also provides a method of manufacturing odorant receptor 6 protein. The method includes the following steps: (a) providing a cell transformed with an isolated DNA comprising a nucleotide sequence that encodes an amino acid sequence of SEQ ID NO: 18; (b) culturing the cell; and (c) collecting from the cell or the medium of the cell the polypeptide encoded by the polynucleotide sequence.

The present invention also provides a method of manufacturing odorant receptor 7 protein. The method includes the following steps: (a) providing a cell transformed with an isolated DNA comprising a nucleotide sequence that encodes an amino acid sequence of SEQ ID NO: 20; (b) culturing the cell; and (c) collecting from the cell or the medium of the cell the polypeptide encoded by the polynucleotide sequence.

The present invention also provides a method for identifying a mosquito olfaction molecule binding compound. The method includes the following steps: (a) providing an isolated mosquito olfaction molecule; (b) contacting a test agent with the isolated mosquito olfaction molecule; and (c) detecting whether the test agent is bound to the isolated mosquito olfaction molecule. Methods of detection are well known in the art. In certain embodiments, the isolated mosquito olfaction molecule further comprises a polypeptide having an amino acid sequence as set forth in SEQ ID NO: 2 or variants thereof as described herein (As used herein this statement means conservatively modified variants, hybridization variants, and variants to which antibodies bind specifically). In alternate embodiments, the isolated mosquito olfaction molecule further comprises a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO. 4, SEQ ID NO. 6, SEQ ID NO. 8, SEQ ID NO. 14, SEQ ID NO. 16, SEQ ID NO. 18, SEQ ID NO. 20. conservatively modified SEQ ID NO: 4, conservatively modified SEQ ID NO: 6, conservatively modified SEQ ID NO: 8, conservatively modified SEQ ID NO: 14, conservatively modified SEQ ID NO: 16, conservatively modified SEQ ID NO: 18, and conservatively modified SEQ ID NO: 20. In other embodiments, contacting the test agent with the isolated mosquito olfaction molecule further comprises contacting under native conditions. In alternate embodiments, detecting specific binding of the test agent to the isolated mosquito olfaction molecule further comprises immunoprecipitation.

The present invention also presents a screening method for identifying a compound that inhibits binding of mosquito arrestin to a mosquito odorant receptor. The method includes the following steps: (a) providing an antibody that binds to an isolated mosquito olfaction molecule; (b) providing a mosquito olfaction molecule binding compound; (c) providing a test sample comprising the mosquito arrestin polypeptide and mosquito odorant receptor; (d) combining the mosquito olfaction molecule binding compound, the antibody, and the test sample in reaction conditions that allow a complex to form in the absence of the mosquito olfaction molecule binding compound., wherein the complex includes the antibody, mosquito arrestin and mosquito odorant receptor; and (e) determining whether the mosquito olfaction molecule binding compound decreases the formation of the complex, wherein a decrease indicates that the mosquito olfaction molecule binding compound is a compound that inhibits the binding of mosquito arrestin to mosquito odorant receptor. In certain embodiments, the mosquito odorant receptor further comprises a polypeptide having any of the following sequences: SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, conservatively modified SEQ ID NO: 4, conservatively modified SEQ ID NO: 6, conservatively modified SEQ ID NO: 8, conservatively modified SEQ ID NO: 16, conservatively modified SEQ ID NO: 18, conservatively modified SEQ ID NO: 20 or conservatively modified SEQ ID NO: 14.

Various features and advantages of the invention will be apparent from the following detailed description and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is the nucleotide sequence (SEQ ID NO: 1) of arrestin 1 isolated from *Anopheles gambiae.*

FIG. 2 is the deduced amino acid sequence of arrestin 1 isolated from *Anopheles gambiae* (SEQ ID NO: 2).

FIG. 3a-b are the nucleotide sequence (SEQ ID NO: 9) and deduced amino acid sequence (SEQ ID NO: 4) of odorant receptor 1 isolated from *Anopheles gambiae.*

FIG. 4a-b are the nucleotide sequence (SEQ ID NO: 10) and deduced amino acid sequence (SEQ ID NO: 6) of odorant receptor 2 isolated from *Anopheles gambiae.*

FIG. 5a-b are the nucleotide sequence (SEQ ID NO: 11) and deduced amino acid sequence (SEQ ID NO: 8) of odorant receptor 3 isolated from *Anopheles gambiae.*

FIG. 6a-b are the nucleotide sequence (SEQ ID NO: 12) and deduced amino acid sequence (SEQ ID NO: 14) of odorant receptor 4 isolated from *Anopheles gambiae.*

FIG. 7 is a table of preferred codons used to deduce amino acid sequences from nucleotide sequences for *Anopheles gambiae.*

FIG. 8 is a table listing cDNA and polypeptide sequences with corresponding SEQ ID numbers and Figure numbers.

FIG. 9a-b are the nucleotide sequence (SEQ ID NO: 21) and deduced amino acid sequence (SEQ ID NO: 16) of odorant receptor 5 isolated from *Anopheles gambiae.*

FIG. 10a-b are the nucleotide sequence (SEQ ID NO: 22) and deduced amino acid sequence (SEQ ID NO: 18) of odorant receptor 6 isolated from *Anopheles gambiae.*

FIG. 11a-b are the nucleotide sequence (SEQ ID NO: 23) and deduced amino acid sequence (SEQ ID NO: 20) of odorant receptor 7 isolated from *Anopheles gambiae.*

DETAILED DESCRIPTION OF THE INVENTION

Arrestins interact with odorant receptors to cause changes in cellular function. Interruption of normal arrestin function will lead to over stimulation of the olfaction system. Consequently, substances that block the arrestin - odorant receptor interaction can interfere with a mosquito's ability to home in on sources of bloodmeal, such as humans. Screening for substances that modulate arrestin - odorant receptor interaction is therefore useful for identifying pest control agents and for treatment of malaria. The deduced amino acid sequence and arrestin contains several domains implicated in arrestin function. The motifs potention consensus Src homology 3 (SH3) binding sites. Cohen, *et al.*, 1995, Cell, 80:237. Sequence comparisons with the DDBJ/EMBL/GenBank and SWISSPROT databases were performed using the GCG software. Devereux, *et al.*, 1984, Nucleic Acids Res., 12:387-395. Protein alignment was also performed using the Clustal W software package. Thompson, *et al.*, 1994, Nucleic Acids Res, 22:4673-4680. Additionally,

arrestin has been submitted to the GenBank database with accession No. AY017417.

As used herein, "native conditions" means natural conditions as found within the ordinary conditions found within *Anopheles gambiae.*

As used herein, "stringent conditions" means the following: hybridization at 42° C in the presence of 50% formamide; a first wash at 65° C with about 2 x SSC containing 1% SDS; followed by a second wash at 65 ° C with 0.1 x SSC. Salt concentrations and temperature may be modified. Such modifications may be found in Sambrook *et al.*, 1989, Molecular Cloning: A Laboratory Manual (2nd Edition), Cold Spring Harbor Press, Cold Spring Harbor, N.Y.The hybridizing part of the nucleic acid is generally at least 15 nucleotides in length.

As used herein, "purified polypeptide" means a polypeptide that is substantially free from compounds normally associated with the polypeptide in the natural state. The absence of such compounds may be determined by detection of protein bands subsequent to SDS-PAGE. Purity may also be assessed in other ways known to those of ordinary skill in the art. The term, as defined herein, is not intended to exclude (1) synthetic or artificial combinations of the polypeptides with other compounds, (2) polypeptides having minor impurities which do not interfere with biological activity.

As used herein, "isolated polynucleotide" means a polynucleotide having a structure that is not identical to any naturally occurring nucleic acid or of any

fragment of a naturally occurring genomic nucleic acid spanning more than three separate genes. Thus, the term includes (1) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (2) a separate molecule of a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (3) a recombinant nucleotide sequence that is part of a gene encoding a fusion protein. This definition of "isolated polynucleotide" supersedes and controls all other definitions known in the art.

As used herein, "hybridization probe" means nucleic acid that is labeled for detection, such as labeling with radiation. Hybridization probes are well known in the art.

As used herein, "culturing the cell" means providing culture conditions that are conducive to polypeptide expression. Such culturing conditions are well known in the art.

As used herein, "operably linked" means incorporated into a genetic construct so that expression control sequences effectively control expression of a gene of interest.

As used herein, "protein" means any peptide-linked chain of amino acids, regardless of length or post-translational modification, *e.g.*, glycosylation or phosphorylation.

As used herein, "sequence identity" means the percentage of identical subunits at corresponding positions in two sequences when the two sequences are aligned to maximize subunit matching, i.e., taking into account gaps and insertions. When a subunit position in both of the two sequences is occupied by the same monomeric subunit, *e.g.*, if a given position is occupied by an adenine in each of two DNA molecules, then the molecules are identical at that position. For example, if 7 positions in a sequence 10 nucleotides in length are identical to the corresponding positions in a second 10-nucleotide sequence, then the two sequences have 70% sequence identity. Preferably, the length of the compared sequences is at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 100 nucleotides. Sequence identity is typically measured using sequence analysis software (*e.g.*, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705).

As used herein, "mosquito olfaction molecule" means a polypeptide that is involved in the modulation of the mosquito olfaction system. By way of illustration, and not limitation, mosquito olfaction molecules have the following characteristics: (1) G protein-coupled seven-transmembrane domain receptors, (2) sequence conservation regarding positions of a subset of introns and the length of the deduced protein, (3) they are selectively expressed in olfactory receptor neurons, and (4) they have highly conserved structural motifs. Odorant receptors 3, 4 and 5 are clustered

tightly together within the *A. gambaie* genome. Odorant receptor 5 and odorant receptor 4 are separated by 310 bp while odorant receptor 4 and odorant receptor 3 are separated by 747 bp. An additional characteristic of odorant and taste receptor genes is the close chromosomal linkage. Such linkage has been demonstrated in the *D. melanogaster* and odorant receptor genes from *C. elegans* and mouse. Clyne, *et al.*, 1999, Neuron, 22:327-338; Vosshall, *et al.*, 1999, Cell, 96:725-736; Vosshall, *et al.*, 2000, Cell, 102:147-159; Clyne, *et al.*, 2000, Science, 287:1830-1834; Gao and Chess 1999, Genomics, 60:31-39; Troemel, *et al.*, 1995, Cell, 83:207-218; Xie, *et al.*, 2000, Genome, 11:1070-1080. Fox *et. al.*, 2001, PNAS 98:14693-14697. This group of molecules includes odorant receptor 1 (SEQ ID NO: 4), odorant receptor 2 (SEQ ID NO: 6), odorant receptor 3 (SEQ ID NO: 8), odorant receptor 4 (SEQ ID NO: 14), odorant receptor 5 (SEQ ID NO: 16), odorant receptor 6 (SEQ ID NO: 18), odorant receptor 7 (SEQ ID NO: 20), arrestin 1 (SEQ ID NO: 2) and variants thereof as described herein.

As used herein, "odorant receptor" means any molecule performing the functional role of an odorant receptor, as described herein and in the scientific literature. Examples of odorant receptors included, but are not limited to, odorant receptor 1, odorant receptor 2, odorant receptor 3, odorant receptor 4, odorant receptor 5, odorant receptor 6, and odorant receptor 7.

As used herein, "mosquito olfaction molecule binding compound" means a compound that specifically binds to a mosquito olfaction molecule. Mosquito

olfaction molecules additionally include polypeptides having the characteristics noted in the definition of the term.

As used herein, "mosquito olfaction molecule-specific antibody" means an antibody that binds to a mosquito olfaction molecule. The term includes polyclonal and monoclonal antibodies.

As used herein, "substantially pure protein" means a protein separated from components that naturally accompany it. Typically, the protein is substantially pure when it is at least 60%, by weight, free from the proteins and other naturally-occurring organic molecules with which it is naturally associated. In certain embodiments, the purity of the preparation is at least 75%, more preferably at least 90%, 95% and most preferably at least 99%, by weight. A substantially pure mosquito olfaction molecule protein can be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid encoding a mosquito olfaction molecule polypeptide, or by chemical synthesis. Purity can be measured by any appropriate method, *e.g.*, column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis. A chemically-synthesized protein or a recombinant protein produced in a cell type other than the cell type in which it naturally occurs is, by definition, substantially free from components that naturally accompany it. Accordingly, substantially pure proteins include those having sequences derived from eukaryotic organisms but synthesized in *E. coli* or other prokaryotes.

As used herein, "fragment", as applied to a polypeptide (*e.g.*, arrestin 1 polypeptide), means at least about 10 amino acids, usually about 20 contiguous amino acids, preferably at least 40 contiguous amino acids, more preferably at least 50 amino acids, and most preferably at least about 60 to 80 or more contiguous amino acids in length. Such peptides can be generated by methods known to those skilled in the art, including proteolytic cleavage of the protein, de novo synthesis of the fragment, or genetic engineering.

As used herein, "test sample" means a sample that contains arrestin 1, or conservatively modified variant thereof, in combination with at least one of the following: odorant receptor 1, odorant receptor 2, odorant receptor 3, odorant receptor 5, odorant receptor 6, odorant receptor 7, odorant receptor 4, conservatively modified variants of the above, or other odorant receptors known in the art.

As used herein, "vector" means a replicable nucleic acid construct, *e.g.*, a plasmid or viral nucleic acid. Preferably, expression is controlled by an expression control sequence.

As used herein, “conservatively modified” applies to both amino acid and nucleic acid sequences. Regarding nucleic acid sequences, conservatively modified refers to those nucleic acids which encode identical or conservatively modified variants of the amino acid sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For example, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine.

Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of ordinary skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine; and UGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide of the present invention is implicit in each described polypeptide sequence and incorporated herein by reference.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Thus, any number of amino acid residues selected from the group of integers consisting of from 1 to 15 can be so altered. Thus, for example, 1, 2, 3, 4, 5, 7, or 10 alterations can be made. Conservatively modified variants typically provide similar biological activity as the unmodified polypeptide sequence from which they are derived. For example, substrate specificity, enzyme activity, or ligand/receptor binding is generally at least 30%, 40%, 50%, 60%, 70%,

80%, or 90% of the native protein for it's native substrate. Conservative substitution tables providing functionally similar amino acids are well known in the art. The following six groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Serine (S), Threonine (T); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W). See also, Creighton (1984) Proteins W.H. Freeman and Company.

As used herein, "immunogenic fragment" means the fragment of a polypeptide that is capable of eliciting an immunogenic response.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present document, including definitions, will control. Unless otherwise indicated, materials, methods, and examples described herein are illustrative only and not intended to be limiting.

Structure and Function

The genes disclosed herein have homology to corresponding arrestin and odorant receptor *Drosophila melanogaster* genes. Fox, *et al.*, 2001, PNAS 98:14693-14697. The genes disclosed herein have the utility disclosed within this patent application.

A full-length *Anopheles gambiae* arrestin 1 cDNA has been cloned and sequenced. The arrestin 1 cDNA clone contains 1964 bp and includes a complete open reading frame that encodes a protein 383 amino acids in length, as seen in Figure 1. The open reading frame from the methionine includes 383 amino acids, yielding a slightly basic polypeptide (PI=8.0) with a predicted molecular weight of 42.8 KD.

A full-length *Anopheles gambiae* odorant receptor 1 genomic DNA has been sequenced. The odorant receptor 1 genomic DNA contains 3895 bp and includes a deduced open reading frame that encodes a protein 394 amino acids in length.

A full-length *Anopheles gambiae* odorant receptor 2 genomic DNA has been sequenced. The odorant receptor 2 genomic DNA contains 4985 bp and includes a deduced open reading frame that encodes a protein 380 amino acids in length.

A full-length *Anopheles gambiae* odorant receptor 3 genomic DNA has been sequenced. The odorant receptor 3 genomic DNA contains 2083 bp and includes a deduced open reading frame that encodes a protein 411 amino acids in length.

A full-length *Anopheles gambiae* odorant receptor 4 genomic DNA has been sequenced. The odorant receptor 4 genomic DNA contains 2374 bp and includes a deduced open reading frame that encodes a protein 394 amino acids in length.

A full-length *Anopheles gambiae* odorant receptor 5 genomic DNA has been sequenced. The odorant receptor 5 genomic DNA contains 2272 bp and includes a deduced open reading frame that encodes a protein 391 amino acids in length.

A partial *Anopheles gambiae* odorant receptor 6 genomic DNA has been sequenced. The odorant receptor 6 genomic DNA contains 931 bp and includes a deduced open reading frame that encodes a protein 157 amino acids in length.

A full-length *Anopheles gambiae* odorant receptor 7 genomic DNA has been sequenced. The odorant receptor 7 genomic DNA contains 11,103 bp and includes a deduced open reading frame that encodes a protein 401 amino acids in length.

Expression Control Sequences and Vectors

The mosquito olfaction molecules of this invention can be used in a method to identify a mosquito olfaction molecule binding compound. If desired, the mosquito olfaction molecule binding compounds may be further tested for ability to inhibit binding of arrestin to an odorant receptor. Methods for this test are described herein. In certain embodiments, the DNA that encodes the arrestin 1 polypeptide ("ARR1 DNA") may be cloned into an expression vector, i.e., a vector wherein ARR1

DNA is operably linked to expression control sequences. The need for expression control sequences will vary according to the type of cell in which the ARR1 DNA is to be expressed. Generally, expression control sequences include a transcriptional promoter, enhancer, suitable mRNA ribosomal binding sites, and sequences that terminate transcription and translation. One of ordinary skill in the art can select proper expression control sequences. Standard methods can be used by one skilled in the art to construct expression vectors. See generally, Sambrook *et al.*, 1989, Molecular Cloning: A Laboratory Manual (2nd Edition), Cold Spring Harbor Press, Cold Spring Harbor, N.Y. Vectors useful in this invention include, but are not limited to plasmid vectors and viral vectors.

All other nucleic acid sequences disclosed herein may also be operably linked to expression control sequences. The expression control sequences described above may be used. As mentioned above, methods known to those of ordinary skill in the art may be used to insert nucleic acid sequences into expression control sequences. Methods known to those of ordinary skill in the art may be used to introduce the nucleic acid and expression control sequence into eukaryotic and/or prokaryotic cells. An example of prokaryotic cells is BL21 (DE3)pLysS bacteria. An example of eukaryotic cells is Sf9.

In certain embodiments of the invention, ARR1 DNA is introduced into, and expressed in, a prokaryotic cell, *e.g.*, BL21 (DE3)pLysS bacteria.

In certain embodiments of the invention, the ARR1 DNA is introduced into, and expressed in, a eukaryotic cell *in vitro*. Eukaryotic cells useful for expressing ARR1 DNA *in vitro* include, but are not limited to Sf9 cells. Transfection of the eukaryotic cell can be transient or stable.

Mosquito Olfaction Molecule-Specific Antibody

An animal is immunized with a mosquito olfaction molecule (*e.g.*, arrestin 1 polypeptide). The animal produces antibodies to the mosquito olfaction molecule. The production and collection of the polyclonal antibodies was performed by Lampire Biological Laboratories, Inc. of Pipersville, PA 18947, using techniques known in the art.

Mosquito Olfaction Molecule Antibody Label

In some embodiments of the invention, the mosquito olfaction molecule-specific antibody includes a detectable label. Many detectable labels can be linked to, or incorporated into, an antibody of this invention. The following are examples of useful labels: radioactive, non-radioactive isotopic, fluorescent, chemiluminescent, paramagnetic, enzyme, or colorimetric.

Examples of useful enzyme labels include malate hydrogenase, staphylococcal dehydrogenase, delta-5-steroid isomerase, alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, and glucoamylase, acetylcholinesterase. Examples of useful radioisotopic labels include ^{3}H, ^{131}I, ^{125}I, ^{32}P, ^{35}S, and ^{14}C. Examples of useful fluorescent labels include fluorescein, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, and fluorescamine. Examples of useful chemiluminescent label types include luminal, isoluminal, aromatic acridinium ester, imidazole, acridinium salt, oxalate ester, luciferin, luciferase, and aequorin.

Antibody labels can be coupled to, or incorporated into antibodies by use of common techniques known to those of ordinary skill in the art. Typical techniques are described by Kennedy *et al.*, 1976, Clin. Chim. Acta, 70:1-31; and Schurs *et al.*, 1977, Clin. Chim. Acta, 81: 1-40. Useful chemical coupling methods include those that use glutaraldehyde, periodate, dimaleimide and m-maleimido-benzyl-N-hydroxy-succinimide ester.

Screening assays

The present invention provides, in part, a screen for mosquito olfaction molecule binding compounds with the ability to interrupt the interaction of arrestin with an odorant receptor. Identifying that a test agent will bind a mosquito olfaction molecule is one part. Once a test agent has demonstrated its ability to bind a mosquito olfaction molecule, it is properly called a mosquito olfaction molecule binding compound. Since it is possible for a mosquito olfaction molecule binding compound to bind without necessarily interrupting the arrestin-odorant receptor interaction, it is proper to further assay in order to determine that the interaction is disrupted. The ability of the mosquito olfaction molecule binding compound to interrupt the arrestin-odorant receptor interaction may be assayed.

In certain embodiments, a test agent is identified as a mosquito olfaction molecule binding compound by the following method. One of the mosquito olfaction molecules is immobilized (*e.g.*, arrestin 1). Polypeptides can be immobilized using methods known in the art. Such methods include the use of Affigel (Biorad) or activated agarose or sepharose to which significant amounts of polypeptides can be directly coupled. The immobilized polypeptide (*e.g.*, arrestin 1) is contacted with the test agent. Unbound test agent can be removed by washing with binding buffer. Then, the bound test agent is eluted by a salt gradient. The material that is bound to the immobilized polypeptide may be purified by SDS-PAGE. Other methods known by one of ordinary skill in the art for identifying an interaction between two

proteins include affinity purification, co-immunoprecipitation, and far-western blotting.

In certain embodiments, the following method is used to screen for substances capable of interrupting arrestin-odorant receptor interaction. The following method of detecting protein-protein interaction will also provide information regarding the lack of protein-protein interactions. The two-hybrid method is a well known genetic assay used to detect protein-protein interactions *in vivo*. See, *e.g.*, Bartel *et al.*, 1993, In Cellular Interactions in Development: A Practical Approach, Oxford University Press, Oxford, pp. 153-179; Chien *et al.*, 1991, Proc. Natl. Acad. Sci. USA, 88:9578-9582; Fields *et al.*, 1989, Nature, 340:245-247; Fritz *et al.*, 1992, Curr. Biol., 2:403-405; Guarente, L., 1993, Proc. Natl. Acad. Sci. USA, 90:1639-1641. There are multiple combinations available between arrestin and the seven odorant receptors. A GAL4 binding domain is linked to an arrestin fragment (*e.g.*, arrestin 1 polypeptide) and a GAL4 transactivation domain is linked to an odorant receptor fragment (*e.g.*, odorant receptor 1 polypeptide). A GAL4 binding site is linked to a reporter gene such as lacZ. All three elements are contacted in the presence and absence of a mosquito olfaction molecule binding compound. The level of expression of the reporter gene is monitored. A decrease in the level of expression of lacZ means that the mosquito olfaction molecule binding compound interrupts the interaction of arrestin with the odorant receptor.

In an alternate embodiment, the following is a method that will identify whether a mosquito olfaction molecule binding compound will interrupt the interaction between arrestin and an odorant receptor. The following method of co-immunoprecipitation may make use of the available panel of antibodies to any arrestin or odorant receptor. Since this method makes use of antibodies that demonstrate the ability to immunoprecipitate the mosquito olfaction molecule and other proteins to which it is bound, the ability of a mosquito olfaction molecule binding compound to inhibit the interaction of the mosquito olfaction molecule will serve as the measure of the compound's interruption ability.

Also disclosed herein is a method of modulating arrestin 1 biological activity. In certain embodiments, the method comprises administering an arrestin 1 biological activity-modulating amount of a mosquito olfaction molecule binding compound. Upon administration, arrestin 1 is contacted with the mosquito olfaction molecule binding compound. Such contact results in modulating arrestin 1 biological activity. The mosquito olfaction molecule binding compound may be administered as an aerosol, solid, or liquid, such that delivery occurs through contact with the body of the target subject. For example, administration may occur by absorption through the exterior surfaces of the target subject, ie. mosquitoes, or by intake through other apertures of the target subject [proboscis (or other feeding aperture), or spiracles (or other respiratory apertures]. An activity-modulating amount of mosquito olfaction molecule binding compound is an amount that is

sufficient to prohibit at least about 50% of the arrestin 1 (SEQ ID NO: 2) molecules from interacting with any odorant receptors.

All citations and references described in this patent application are hereby incorporated herein by reference, in their entirety. Also incorporated in this specification are the exhibits filed herewith. The present invention is further illustrated by the following specific examples. The examples are provided for illustration only and are not to be construed as limiting the scope or content of the invention in any way.

Example 1

Protein expression

A cDNA encoding arrestin 1 is subcloned into the pBlueScript II (KS) vector (Novagen, Madison, WI) at the BamHI/NdeI restriction sites for DNA sequencing. The cDNA encoding arrestin 1 is subsequently subcloned into the bacterial expression plasmid pET15b (Novagen, Madison, WI). The bacterial expression plasmid containing the arrestin 1 cDNA is transformed into BL21 (DE3)pLysS bacteria (Novagen, Madison, WI) for high levels of arrestin 1 expression. Methods are known in the art for isolating the expressed protein.

Expression of other nucleic acids disclosed herein is achieved by using the above-referenced method. Once the odorant receptor is in protein form, it may be used as described within this application.

Example 2

Mosquito Olfaction Molecule Specific Antibody

The cDNA encoding arrestin 1 is subcloned into the bacterial expression plasmid pET15b (Novagen, Madison, WI). The vector is transformed into BL21 (DE3)pLysS bacteria (Novagen, Madison, WI) for high levels of arrestin 1 expression. Rapid purification is performed using His-Bind affinity Resin (Novagen, Madison, WI). Native recombinant arrestin 1 is then denatured using gel purification on SDS-polyacrylamide gel electrophoresis followed by staining with 0.05% Coomassie Brilliant Blue (Sigma-Aldrich, St. Louis, MO). Polyclonal antibodies were generated in rabbits by Lampire Biological Laboratories, Inc. of Pipersville, PA 18947. Polyclonal antibodies may be generated for any of the odorant receptors disclosed herein.

Example 3

Identification of a mosquito olfaction molecule binding compound

Arrestin 1 polypeptide is expressed in and purified from BL21 (DE3)pLysS bacteria (Novagen, Madison, WI). Arrestin 1 is incubated with a test agent in Phosphate Buffered Saline (pH 7.5), 0.1% Tween-20, and 0.1% broad spectrum protease inhibitors for 90 minutes at 4° C. Anti-arrestin 1 polyclonal sera is added to the reaction at a dilution of 1:2000 and incubated for an additional 60 minutes. The complexes, consisting of either polypeptide-antibody or test agent-polypeptide-antibody are isolated by the addition of 1x 10^7 Dynalbeads M280 (sheep anti-Rabbit

IgG) followed by incubation at the same temperature for an additional 60 minutes. Isolation of the complexes is completed by using the DYNAL Magnetic Particle Concentrator (Dynal Inc., Lake Success, NY). The complexes are washed three times with broad spectrum protease inhibitors. Content of the complexes is assayed by SDS-PAGE followed by silver staining and western blotting. Common methods are known by those of ordinary skill in the art for silver staining and western blotting. See generally, Sambrook *et al.*, 2001, Molecular Cloning: A Laboratory Manual (3rd Edition), Cold Spring Harbor Press, Cold Spring Harbor, N.Y. Obviously, the presence of the test agent, polypeptide, and antibody indicates that the test agent binds to the polypeptide.

Example 4

Identification of a compound that inhibits binding of arrestin to an odorant receptor

Arrestin 1 polypeptide and odorant receptor 1 polypeptide are expressed in and purified from BL21 (DE3)pLysS bacteria (Novagen, Madison, WI). Arrestin 1 polypeptide and odorant receptor 1 polypeptide are incubated with a mosquito olfaction molecule binding compound in Phosphate Buffered Saline (pH 7.5), 0.1% Tween-20, and 0.1% broad spectrum protease inhibitors for 90 minutes at 4° C. Anti-arrestin 1 polyclonal sera is added to the reaction at a dilution of 1:2000 and incubated for an additional 60 minutes. The complexes, consisting of either

antibody-arrestin 1-odorant receptor 1 or antibody-arrestin 1, are isolated by the addition of 1x 10 ^{7} Dynalbeads M280 (sheep anti-Rabbit IgG) followed by incubation at the same temperature for an additional 60 minutes (Dynal Inc., Lake Success, NY). Once the isolation of the complexes is completed by using the DYNAL Magnetic Particle Concentrator, (Dynal Inc., Lake Success, NY), the complexes are washed three times with broad spectrum protease inhibitors. The content of the complexes is assayed by SDS-PAGE followed by silver staining and western blotting. Common methods are known by those of ordinary skill in the art for silver staining and western blotting. See generally, Sambrook *et al.*, 2001, Molecular Cloning: A Laboratory Manual (3rd Edition), Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

Example 5

Far western blotting to analyze components of a protein mixture

The protein sample is fractionated on an SDS-PAGE gel. After electrophoresis at a voltage and time that is known in the art, the proteins are transferred from the gels onto a solid support membrane by electroblotting. Transferred membranes may be stained with Ponceau S to facilitate location and identification of specific proteins. Nonspecific sites on the membranes are blocked with standard blocking reagents, and the membranes are then incubated with a

radiolabeled non-antibody protein probe. After washing, proteins that bind to the probe are detected by autoradiography.

The content of the solutions used within this protocol are disclosed in Wiley's Current Protocols in Cell Biology.

The protein sample to be analyzed is resuspended in 1x SDS sample buffer. Approximately 50 to 100 ug can be loaded in each lane of the gel. The samples are separated with SDS-PAGE. The proteins are transferred to nitrocellulose by electroblotting.

After transfer, stain the membrane for 5 min in ~100 ml freshly diluted 1x Ponceau S staining solution. The membrane is then destained by washing it in several changes of deionized water until the proteins are clearly visible. Continue to destain for an additional 5 min in water until the red staining fades.

The membrane is then blocked for 2 hr in 200 ml blocking buffer I at room temperature with gentle agitation. Incubate the membrane in 200 ml of blocking buffer II for 2 hours and rinse the membrane briefly in 100 ml of 1 x PBS.

Prior to probing, the membrane is preincubated for 10 min in 50 ml of 1x probe dilution buffer without the probe at room temperature. The probe is added to the membrane and incubated for 2 hours at room temperature. The membrane is washed with 200 ml 1x PBS for 5 min, room temperature. Repeat the wash step three additional times. Air dry the filter and expose to x-ray film with intensifying screen. An overnight exposure is typically sufficient.

Sequence ID Listing

SEQ ID NO:1
cDNA Nucleic Acid Sequence
1964 nucleotides
Mosquito arrestin 1

ACAGGAACGACGGTTGTGATCCCTCCACTGGTGGTGACACGAATCATAAGCAT
TATTTCATACCTAAAAAACAAAATCTACAAAAAAAAGCTTCATTCCCATCGAAA
AAACTTTCTTGTGAAATCAACCGAGCTAACAAACAACATCCTGTGCAAAATCTA
GCAGTGAAAGTGTGATATCGTATACCTGTACCTGTAAACCGTTGTGCGCGTGT
GTGCCTTTGTGTATCAATTTTGTGGAAAACAGAAAATACATCAAAATGGTTTAC
AATTTCAAAGTCTTCAAGAAGTGCGCCCCTAATGGAAAGGTTACGCTGTACATG
GGCAAGCGTGACTTTGTAGACCACGTTTCCGGCGTTGAACCGATCGATGGTAT
CGTCGTCCTCGATGATGAGTACATTCGTGACAACCGTAAGGTATTCGGTCAGAT
TGTCTGCAGTTTCCGCTACGGCCGCGAAGAGGACGAGGTGATGGGACTAAACT
TCCAGAAGGAGTTATGCCTCGCTTCCGAACAGATCTACCCGCGTCCGGAAAAG
TCGGACAAGGAGCAGACCAAGCTCCAGGAGCGACTGCTGAAGAAGCTGGGTTC
GAACGCCATCCCGTTCACGTTCAACATCTCGCCGAATGCTCCGTCTTCGGTCAC
GCTGCAGCAGGGCGAAGATGATAATGGAGACCCGTGCGGTGTGTCGTACTACG
TGAAGATCTTTGCCGGTGAGTCGGAAACCGATCGTACGCACCGTCGCAGCACC
GTTACGCTCGGCATACGCAAGATCCAGTTCGCACCGACCAAGCAGGGCCAGCA
GCCGTGCACGCTGGTGCGCAAGGACTTTATGCTAAGCCCGGGAGAGCTGGAGC
TCGAGGTCACACTAGACAAGCAGCTGTACCTGCACGGGGAGCGAATAGGCGTC
AACATCTGCATCCGCAACAACTCGAACAAAATGGTCAAGAAGATTAAGGCCAT
GGTCCAGCAGGGTGTGGATGTGGTGCTGTTCCAGAATGGTAGCTACCGCAACA
CAGTGGCATCGCTGGAGACTAGCGAGGGTTGCCCAATTCAGCCCGGCTCCAGT
CTGCAGAAGGTAATGTACCTCACGCCGCTGCTGTCCTCGAACAAGCAGCGACG
TGGCATCGCCCTGGACGGTCAGATCAAGCGTCAGGATCAGTGTTTGGCCTCGA
CAACCCTCTTGGCTCAACCGGATCAGCGAGATGCTTTCGGCGTTATCATATCGT
ATGCCGTAAAGGTTAAGCTTTTCCTCGGCGCACTCGGCGGCGAGCTGTCGGCG
GAACTTCCATTTGTGCTGATGCACCCAAAGCCCGGCACCAAGGCTAAGGTCAT
CCATGCCGACAGCCAGGCCGACGTAGAAACTTTCCGACAGGATACAATCGACC
AGCAGGCATCAGTTGACTTTGAATAGACGACGCAACGGTTTGGAAATGCTACC
TACTACCCCAGGCATGGGCTAACACGACGAACGAACTACTACTACTAAGCATA
AAAAACAGGAAAAAAAATGGAAAACTTAAAAAATGGATCATACAACCGAACGC
AAACGACCTACGACGATCGATCTCACTTCCCCGTCTTTTTCATCCTAAGCAATA
GAACGATGGTAGAAAAGGAAGATAAAGATGGAGAGAAAGTCACGTGTATCAAT
GACGACGACTACCAAAACTGAAGACGTAACACATGTTCCCCAGCGAGCGGTAA
CTGTTCTGTTCTGACACCTTCCGCTCGACAATGTACCTTTTAAAAACATACAAA

TTAGAAGTCGTCTTCACTACCTTCAACCAATCCAGCCACTTTGGTATATACTTTT
CATAGAATCCTTCTGAGCGCAAGGACCCTATTGAAATTCAGTGTTATTTTGTAA
CTGCGACCAAATGCCTAGCTGAATGTTGTTGAACGAGTTATGTACATCAAAAGA
TTGAATAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO:2
Amino Acid Sequence
383 residues
Mosquito arrestin 1

MVYNFKVFKKCAPNGKVTLYMGKRDFVDHVSGVEPIDGIVVLDDEYIRDNRKVF
GQIVCSFRYGREEDEVMGLNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKLG
SNAIPFTFNISPNAPSSVTLQQGEDDNGDPCGVSYYVKIFAGESETDRTHRRSTVT
LGIRKIQFAPTKQGQQPCTLVRKDFMLSPGELELEVTLDKQLYLHGERIGVNICIR
NNSNKMVKKIKAMVQQGVDVVLFQNGSYRNTVASLETSEGCPIQPGSSLQKVMY
LTPLLSSNKQRRGIALDGQIKRQDQCLASTTLLAQPDQRDAFGVIISYAVKVKLFL
GALGGELSAELPFVLMHPKPGTKAKVIHADSQADVETFRQDTIDQQASVDFE

SEQ ID NO:3
cDNA Nucleic Acid Sequence
1239 nucleotides
Mosquito odorant receptor 1

ATGAAGCTGAACAAACTGAACCCACGGTGGGATGCGTACGATCGACGGGATTCGTTCTGGTTGCAGTTGCTTTGTTT
GAAATATTTAGGCCTATGGCCACCGGAAGATACGGATCAGGCAACGCGGAACCGGTACATCGCGTACGGTTGGGCTT
TGCGGATCATGTTTCTACATCTGTACGCTCTAACGCAAGCCCTATACTTCAAGGATGTGAAGGATATTAATGACATC
GCAAATGCATTGTTCGTGCTTATGACTCAAGTGACGTTGATCTACAAGCTGGAAAAGTTTAACTACAACATCGCACG
GATTCAGGCTTGTCTGCGCAAGCTTAACTGCACACTGTATCACCCGAAACAGCGCGAAGAATTCAGCCCCGTTTTAC
AATCGATGAGTGGAGTGTTTTGGCTGATGATCTTTCTCATGTTTGTGGCTATCTTCACCATCATCATGTGGGTTATG
TCGCCAGCCTTCGACAATGAACGTCGTCTGCCCGTGCCGGCCTGGTTCCCGGTGGACTATCACCATTCGGACATAGT
GTACGGTGTACTGTTCCTGTATCAAACCATTGGAATCGTCATGAGCGCAACGTACAACTTCTCGACCGATACCATGT
TTTCCGGCTTGATGCTACACATAAATGGACAAATTGTGCGGCTTGGTAGTATGGTTAAAAAGCTTGGACATGACGTC
CCTCCCGAACGCCAATTGGTCGCAACGGATGCGGAATGGAAAGAGATGCGAAAGCGCATCGACCATCACTCCAAAGT
GTACGGTACGATGTACGCTAAAGTAACGGAGTGTGTGCTGTTTCACAAGGACATCTTAAGGATCTATCTTCGCGCAA
GTATGCGCGTCTGTAATTATCATTTGTATGACACTGCTGCAACTACCGGGGGCGATGTTACGATGGCCGATCTGCTG
GGCTGTGGGGTCTATTTGCTAGTAAAGACATCGCAAGTGTTTATTTTCTGTTACGTAGGGAATGAAATCTCCTATAC
GACGGATAAATTTACAGAGTTTGTTGGGTTTTCCAACTACTTCAAGTTCGATAAGCGTACCAGCCAAGCAATGATAT
TTTTTCTGCAAATGACTCTTAAAGATGTTCACATCAAGGTGGGAAGTGTCTTGAAGGTTACGCTAAATCTTCACACA
TTTTTGCAGATTATGAAGCTATCGTACTCCTATCTGGCCGTACTTCAGAGCATGGAATCAGAGTAATGGTGTTAATA
TCCT**TAA**

SEQ ID NO:4
Amino Acid Sequence
394 residues
Mosquito odorant receptor 1

MKKDSFFKMLNKHRWILCLWPPEDTDQATRNRYIAYGWALRIMFLHLYALTQA
LYFKDVKDINDIANALFVLMTQVTLIYKLEKFNYNIARIQACLRKLNCTLYHPKQ
REEFSPVLQSMSGVFWLMIFLMFVAIFTIIMWVMSPAFDNERRLPVPAWFPVDY
HHSDIVYGVLFLYQTIGIVMSATYNFSTDTMFSGLMLHINGQIVRLGSMVKKLG
HDVPPERQLVATDAEWKEMRKRIDHHSKVYGTMYAKVTECVLFHKDILRIYLR
ASMRVCNYHLYDTAATTGGDVTMADLLGCGVYLLVKTSQVFIFCYVGNEISYTD
KFTEFVGFSNYFKFDKRTSQAMIFFLQMTLKDVHIKVGSVLKVTLNLHTFLQIM
KLSYSYLAVLQSMESEZ

SEQ ID NO:5
cDNA Nucleic Acid Sequence
1142 nucleotides
Mosquito odorant receptor 2

ATGCTGATCGAAGAGTGTCCGATAATTGGTGTCAATGTGCGAGTGTGGCTGTTCTGGTCGTATCTGCGGCGGCCGCG
GTTGTCCCGCTTTCTGGTCGGCTGCATCCCGGTCGCCGTGCTGAACGTTTTCCAGTTCCTGAAGCTGTACTCGTCCT
GGGGCGACATGAGCGAGCTCATCATCAACGGATACTTTACCGTGCTGTACTTTAACCTCGTCCTCCGAACCTCCTTT
CTCGTGATCAATCGACGGAAATTTGAGACATTTTTTGAAGGCGTTGCCGCCGAGTACGCTCTCCTCGAGAAAAATGA
CGACATCCGACCCGTGCTGGAGCGGTACACACGGCGGGGACGCATGCTATCGATATCGAATCTGTGGCTCGGCGCCT
TCATTAGTGCCTGCTTTGTGACCTATCCTCTGTTTGTGCCCGGGCGCGGCCTACCGTACGGCGTCACGATACCGGGC
GTGGACGTGCTGGCCACCCCGACCTACCAGGTCGTGTTTGTGCTGCAGGTTTACCTTACCTTCCCCGCCTGCTGCAT
GTACATCCCGTTCACCAGCTTCTACGCGACCTGCACGCTGTTTGCGCTCGTCCAGATAGCGGCCCTAAAGCAACGGC
TCGGACGCTTGGGGCGCCACAGCGGCACGATGGCTTCGACCGGACACAGCGCCGGCACACTGTTCGCCGAGCTGAAG
GAGTGTCTAAAGTATCACAAACAAATCATCCAATATGTTCATGATCTCAACTCACTCGTCACCCATCTGTGTCTGCT
GGAGTTCCTGTCGTTCGGGATGATGCTGTGCGCACTGCTGTTTCTGCTAAGCATTAGCAATCAGCTGGCACAGATGA
TAATGATTGGATCGTACATCTTCATGATACTCTCGCAGATGTTTGCCTTCTATTGGCATGCGAACGAGGTACTGGAG
CAGAGCCTAGGCATTGGCGATGCCATTTACAATGGAGCGTGGCCGGACTTTGAGGAACCGATAAGGAAACGGTTGAT
TCTAATTATTGCACGTGCTCAGCGACCGATGGTGGTAAGATTAAAGTCGGCAACGTGTACCCGATGACGTTGGAAAT
GTTTCAAAAATTGCTCAACGTGTCCTACTCCTATTTCACACTGCTGCGCCGAGTGTACAAC**TAA**

SEQ ID NO:6
Amino Acid Sequence
380 residues
Mosquito odorant receptor 2

MLIEECPIIGVNVRVWLFWSYLRRPRLSRFLVGCIPVAVLNVFQFLKLYSSWGDM
SELIINGYFTVLYFNLVLRTSFLVINRRKFETFFEGVAAEYALLEKNDDIRPVLER

YTRRGRMLSISNLWLGAFISACFVTYPLFVPGRGLPYGVTIPGVDVLATPTYQVV
FVLQVYLTFPACCMYIPFTSFYATCTLFALVQIAALKQRLGRLGRHSGTMASTGH
SAGTLFAELKECLKYHKQIIQYVHDLNSLVTHLCLLEFLSFGMMLCALLFLLSIS
NQLAQMIMIGSYIFMILSQMFAFYWHANEVLEASLGIGDAIYNGAWPDFEEPIRK
RLILIIARAQPTDGGKIKVGNVYPMTLEMFQKLLNVSYSYFTLLRRVYN

SEQ ID NO:7
cDNA Nucleic Acid Sequence
1236 nucleotides
Mosquito odorant receptor 3

ATGCCTTCTGAGCGGCTTCGTCTCATTACTTCCTTCGGAACTCCTCAAGACAAACGCACGATGGTACTGCCAAAATT
AAAGGATGAAACAGCAGTGATGCCGTTTCTGCTGCAAATTCAAACCATTGCCGGACTGTGGGGTGACCGTTCCCAGC
GGTACCGTTTTTATCTCATCTTTTCCTACTTCTGCGCGATGGTGGTTCTACCCAAAGTGCTGTTCGGTTATCCAGAT
CTCGAGGTTGCGGTACGCGGCACGGCCGAGCTGATGTTCGAATCGAACGCATTCTTCGGCATGCTAATGTTTTCCTT
TCAACGCGACAACTACGAGCGATTGGTGCATCAGCTGCAGGATCTGGCAGCTCTAGTCCTCCAAGACCTACCCACAG
AGCTGGGAGAGTACCTGATCTCAGTGAACCGACGGGTCGATCGGTTCTCCAAAATTTACTGCTGCTGTCACTTTTCC
ATGGCAACGTTCTTTTGGTTCATGCCCGTCTGGACGACCTATTCCGCCTACTTTGCTGTGCGCAACAGCACGGAACC
GGTCGAGCACGTGTTGCACCTCGAGGAAGAGCTGTACTTCCTGAACATTCGGACTTCGATGGCGCACTATACGTTTT
ATGTGGCCATTATGTGGCCCACGATCTATACGCTCGGGTTTACCGGTGGCACAAAGCTGCTGACCATTTTCAGCAAT
GTTAAGTACTGTTCGGCCATGCTGAAGCTCGTTGCACTCCGAATCCACTGTCTAGCGAGAGTAGCGCAAGACCGAGC
GGAAAAGGAGCTGAACGAGATTATTTCCATGCATCAGCGGGTACTCAACTGCGTGTTCCTGCTGGAGACGACATTCC
GCTGGGTATTTTTCGTGCAGTTCATTCAGTGTACAATGATCTGGTGCAGTCTCATCCTCTACATAGCGGTGACGGGG
TTCAGCTCGACGGTAGCGAATGTATGTGTCCAGATCATTTTGGTGACGGTGGAAACTTACGGCTACGGCTACTTCGG
AACAGATCTAACCACGGAGGTGCTTTGGAGCTATGGCGTTGCCCTCGCCATTTACGATAGCGAGTGGTACAAGTTTT
CCATTTCGATGCGCCGCAAACTTCGACTGCTACTGCAACGATCCCAAAAACCGCTCGGCGTAACGGCGGGAAAGTTT
CGCTTCGTCAATGTGGCCCAGTTTGGCAAGATGCTCAAGATGTCCTATTCATTTTACGTAGTACTGAAGGAGCAGTT
T**TAG**

SEQ ID NO:8
Amino Acid Sequence
411 residues
Mosquito odorant receptor 3

MPSERLRLITSFGTPQDKRTMVLPKLKDETAVMPFLLQIQTIAGLWGDRSQRYR
FYLIFSYFCAMVVLPKVLFGYPDLEVAVRGTAELMFESNAFFGMLMFSFQRDNY
ERLVHQLQDLAALVLQDLPTELGEYLISVNRRVDRFSKIYCCCHFSMATFFWFM
PVWTTYSAYFAVRNSTEPVEHVLHLEEELYFLNIRTSMAHYTFYVAIMWPTIYTL
GFTGGTKLLTIFSNVKYCSAMLKLVALRIHCLARVAQDRAEKELNEIISMHQRVL
NCVFLLETTFRWVFFVQFIQCTMIWCSLILYIAVTGFSSTVANVCVQIILVTVETY
GYGYFGTDLTTEVLWSYGVALAIYDSEWYKFSISMRRKLRLLLQRSQKPLGVTA
GKFRFVNVAQFGKMLKMSYSFYVVLKEQF

SEQ ID NO:9
Genomic Nucleic Acid Sequence
3895 nucleotides
Mosquito odorant receptor 1

AGCTTTGTTCATTTATGTTGAAATCTAGCCCATTTTGTATAGTGCTGAACGACGAAGAACATACGAAAGTACCTCGT
CCGAACACTATCAACATTAATTATACCAAGCTAGAAGAAGATATTTATAGTCAAGCCTCAACATCATAGGAAACTTT
AGCAAAACCATTTAATTTACATGATGATAAGTCCCACCTCTTACCCCAGCACAGGTTTGAGAAGGACGAAAGTATCT
TTACGATAATATTACTCTAAGGTAGTTTTTGAATAAAATAAAAATTTACGTGCAAGTGGTGGCATCGGACATCATTC
GAAAGAATCTACTAAGTCATACACACACCCAAGACGACCGACGTAGTTTCATCTAGAAAAAACGGGTCAGCTCCATC
GAACACGTCAGGACATAACTGCGACATGCGTATGGTCAGTTCCACTAGTGCCAACACTGGTTCCAGGGCACTACCTT
CCGAAGCAGTAGAACCTAATGTATTGGAAATTATTAGGACATACTGCAACATGCATATGGCTAGTTCCGCTGGTACC
AACGATGGCACCAGGACACTATCTGCGGCCTTGTAAAATCACTGTAAAATCTATACAAAAACGGCTTTACCCATACT
TTATCACAAAAACGGCAGGTGAGGGCTGGATTGCTTCAAAGCATTAGAAATATATAATTTCAAAGTCCATAATCTCC
TTAAAAGATAGACAaCAGTAGAGAACACATTTAGTGCTCTTTTCGTTCGAGTTAGTTGCCTTCTCAAGTAAGCGTTT
AATGCTCAATTGTTGTAGATTCGTTGGATGACTCTCGCTACGTGCTATAGTGGTCAATACTTCCAATTAGATTTCAT
AATTAGTTTCCAATTGTCCACGGAAAACCCaCAAAAGAAAAAAAAACTTGTATCTAGGGTGGAATTTTTCGAGAACA
ATTGGACACTTCATATGAAAAAGGACAGCTTTTTCAAAATGTTAAATAAACACCGTTGGATCCTTTgttggatttca
attctccaaattctgcagaataattctgcaaattttacaaaactgctcaaccaccaataattccaattaatcatctg
aacatttaaaactgataattaagatgagtaattgcttcgtcatcacctaagaaatcgattagtttggataaaaagaa
caaattgaaatacaataaagtccctgaattttattcgaataacggcttgaactcatttatttcaaaaacctttgaga
aattcctcgttgaaaattggtctcctatagttctgctaacgggccacttcaaaagcaagaactaacaaaatcataat
tatggtgcaagtaactatcagtaccagtaatcgccattaaaaactttcctcaatttgcggctcgttaccggctaaa
tacagagcagagtaacgggaagtgatcaacgtcgctattagtataacgaggaacgccctccgaaggtgtgttgaagg
acctttcaaattgaaaccaagtactgtttccagttttaaattggatagttataaaatgagccgttcaacgatcggg
catcatttgagtttcatcttcgaggagaaatagatcagtgccactgtttaaccgaaagtaatgaagctgaacaaact
gaacccacggtgggatgcgtacgatcgacgggattcgttctggttgcagttgctttgtttgaaatatttagGCCTAT
GGCCACCGGAAGATACGGATCAGGCAACGCGGAACCGGTACATCGCGTACGGTTGGGCTTTGCGGATCATGTTTCTA
CATCTGTACGCTCTAACGCAAGCCCTATACTTCAAGgATGTGAAGGATATTAATgtgagtctctagttagctattag
tgttccacctgtccataatctgtcttttattgggtagGACATCGCAAATGCATTGTTCGTGCTTATGACTCAAGTGA
CGTTGATCTACAAGCTGGAAAAGTTTAACTACAACATCGCACGGATTCAGGCTTGTCTGCGCAAGCTTAACTGCACA
CTGTATCACCCGAAACAGCGCGAAGAATTCAGgtaagcctgctgggaaatatgactaaaaagagtgctaacaaacga
ctctcctccaaatgtagCCCCGTTTTACAATCGATGAGTGGAGTGTTTTGGCTGATGATCTTTCTCATGTTTGTGGC
TATCTTCACCATCATCATGTGGGTTATGTCGCCAGCCTTCGACAATGAACGTCGTCTGCCcGTGCCGGCCTGGTTCC
CGGTGGACTATCACCATTCGGACATAGTGTACGGTGTACTGTTCCTGTATCAAACCATTGGAATCGTCATGAGCGCA
ACGTACAACTTCTCGACCGATACCATGTTTTCCGGCTTGATGCTACACATAAATGGACAAATTGTGCGGCTTGGTAG
TATGGTTAAAAAGgtgagttacggcgactacttgcctccagtaaggacagggagtttgtttccgttatgatatcatt
ttatcagCTTGGACATGACGTCCCTCCCGAACGCCAATTGGTCGCAACGGATGCGGAATGGAAAGAGATGCGAAAGC
GCATCGACCATCACTCCAAAGTGTACGGTACGATGTACGCTAAAGTAACGGAGTGTGTGCTGTTTCACAAGGACATC
TTAAGgtacgaattgggccaattaattgtgtcatttaaaaagcttgacccaacttttcacagcttcggcgatgaagt
gcaggacattttccaagGATCTATCTTCGCGCAAGTATGCGCGTCTGTAATTATCATTTGTATGACACTGCTGCAAc
TACCGGGGgCGATGTTACGATgGCCGATCTGCTGGGCTGTGGGGTCTATTTGCTAGTAAaGACATCGCAAGTGTTTA
TTTTCTGTTACGTAGGGAATGAAATCTCCTATACGgtaggttggacacgtagaggaattaaatgtttgggaagaata
tcaataccaaatagtatgatgtttcgttacagACGGATAAATTTACAGAGTTTGTTGGGTTTTCCAACTACTTCAAG
TTCGATAAGCGTACCAGCCAAGCAATGATATTTTTTCTGCAAATgtgagatagcggtgtatttgtgcagtcagtaca
ttaaatacgttctctatttcagGACTCTTAAAGATGTTCACATCAAGGTGGGAAGTGTCTTGAAGGTTACGCTAAAT
CTTCACACATTTTTGCAGgtatgtaattatgctgtggtatttagcttgaaataagctacaaactttgaaagtaattt
caatctgttttgtagATTATGAAGCTATCGTACTCCTATCTGGCCGTACTTCAGAGCATGGAATCAGAGTAATGGtG
tTAATATCCtTAATGTTGAAATTATATTTTGTTAGATTTATTGCATAAAGTAaTaTTTAATTTTATACATCAAACGT
AAGCCCGCtaGTTTTCAATTAGCCTTTTCCAAAATTTATCAAATTGATTTCGAATTGATTGCAGAGTTTCAGGAATT
TAATCTGATAGGATATCTTGTTTATCCAATAGAGGTGTGGAAGCGTTCCCAAGCCATTCGTTTGATAGTTTATAGCA

CCGTCGAGCAGTTGATCGCTGTGATCGCTAGGCGCACCTGATTTTATCTTTATCTCGCACCTGTTATGGCAAGGGCG
CTTTTCACACGTTTCACACAATATAATGCACATGTATAATGCATTCTTACTTTAGCATTTTTGTTACATATAATACC
AAAATTATGCATTTTTATTCTCACGCAACGATTAGAGGATGACTTcACAAAGGTCCATCTAGTGGTAGGAGGTATAC
AATTATACCTCTCAAAATCTCACAGCAtAATGAGAAACAAAAGGATACCAAGCATACCCTTTTTTTACTTGACAATT
TCATTTGATTTATGTAATAAAGCACTGCaCGTCGACTTCCTAAAA

SEQ ID NO:10

Genomic Nucleic Acid Sequence
4985 nucleotides
Mosquito odorant receptor 2

GGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTCCCTCACCGTGACGTGCTAGAAATGGTTCAACATACTCGT
CCGGCAGAGCGAAGACGACGAACAGCGGAATGTCCCAGGAAATGTAATGAGATATCACAGCAAGTGAACCCAAACCG
AGCTGTGCGCTTTGTGTTGCGCTTTAAAAATGGCCCTTCCTTCGCCGCATCTGCTTGGTTTCACACGCTTTCCCAGG
AAATCCACTGACCACTGGCCACACATCAACCACCGGAGCGGGAGCCTCAGTGCCCAGCGAAGCATATAATTTGCTCA
AAAAGTCACGGTACTCAATTAATTTGATTATAATCAATTTCGTGGCTTCCAACACACCCTTCTTCCACAATCCATCG
CCGAGTGAGCGAGTATAAAGGTGAAGAAACGTACCTTGCGCTTGCTCACTAACTGAACCGGATTTCAAAAAGGAACA
TAAACCGCAACCCACAGCCGAAAATGCTGATCGAAGAGTGTCCGATAATTGGTGTCAATGTGCGAGTGTGGCTGTTC
TGGTCGTATCTGCGGCGGCCGCGGTTGTCCCGCTTTCTGGTCGGCTGCATCCCGGTCGCCGTGCTGAACGTTTTCCA
GTTCCTGAAGCTGTACTCGTCCTGGGGCGACATGAGCGAGCTCATCATCAACGGATACTTTACCGTGCTGTACTTTA
ACCTCGTCGtacgtgggcgaggggaggggcaataaccttcccacttggtggatattttcataccttttccatgtgtt
tttttattctctgtttgttgccatccagCTCCGAACCTCCTTTCTCGTGATCAATCGACGGAAATTTGAGACATTTT
TTGAAGGCGTTGCCGCCGAGTACGCTCTCCTCGAGgtaagtcattggtttttctagtttttgggggagttgtttaca
ccataaccacccccgacggtaacatttgatcgtcccgcgaaaatgtttgtacagAAAAATGACGACATCCGACCCGT
GCTGGAGCGGTACACACGGCGGGGACGCATGCTATCGATATCGAATCTGTGGCTCGGCGCCTTCATTAGTGCCTGCT
TTGTGACCTATCCTCTGTTTGTGCCCGGGCGCGGCCTACCGTACGGCGTCACGATACCGGGCGTGGACGTGCTGGCC
ACCCCGACCTACCAGGTCGTGTTTGTGCTGCAGGTTTACCTTACCTTCCCCGCCTGCTGCATGTACATCCCGTTCAC
CAGCTTCTACGCGACCTGCACGCTGTTTGCGCTCGTCCAGATAGCGGCCCTAAAGCAACGGCTCGGACGCTTGGGGC
GCCACAGCGGCACGATGGCTTCGACCGGACACAGCGCCGGCACACTGTTCGCCGAGCTGAAGGAGTGTCTAAAGTAT
CACAAACAAATCATCCAGtaagtagacgctagtagactcgaccggattgcccttccctcggggagggaggtttgct
atttcgggatgcggcagcacgcatacacacacaaaccggaagccattaattctcccgttttcatgcccgcacgggcact
gggtcatgtttcacatccttccttcctttccaaacacacacacgcgcgcgtgcacgtacagATATGTTCATGATCTC
AACTCACTCGTCACCCATCTGTGTCTGCTGGAGTTCCTGTCGTTCGGGATGATGCTGTGCGCACTGCTGTTTCTGCT
AAGCATTGtaagtaaaatcgaccgacgtgcggtcgctagtccgtctccggactctcatttcgggactcaatcgttcc
atctctcaatagAGCAATCAGCTGGCACAGATGATAATGATTGGATCGTACATCTTCATGATACTCTCGCAGATGTT
TGCCTTCTATTGGCATGCGAACGAGGTACTGGAGCAGGtaatggcgctgaagctgagtttggttgagcggttcgcta
tagatcggctgtcttacattgttgtgtttctgcatggggatcggttttgtttttcctctccatttcagAGCCTAGGC
ATTGGCGATGCCATTTACAATGGAGCGTGGCCGGACTTTGAGGAACCGATAAGGAAACGGTTGATTCTAATTATTGC
ACGTGCTCAGCGACCGATGGTGGTAAGtttggctgatcgatgctctgttcaatgaacatggcacagaaggctgtgta
aatagctgttcattaataagtttttcagaatgtatcgtttttagttgatttaaacgcattgttctatgcaatggta
gcaacaatagaccgcctttattaatccaagcttcctttaggattgatttttatttaagagaaagataaaccatttt
tagtaaccaatttagttacaggaaccaaaatacagaatttattattattattattattattattattattatta
ttattattattattattattattattattattattattattattattataattattattattattattattatt
attattattattaatattattattattattattattactattattattataattattacttttattattat
tattattattattattattattattattattattattattattattattataattatgattattattattatta
ttattattattattattataacaataataattattattattatttattattaattaattaatttattattatta
attattattattgttattcattattatacattattatcataataataattttattatgattattattattattatta
ttattattattattattattattattattcttattattattattattattattattattaatattatttttaatatt

attattattattattactattcttattataattatttttttttattattattattattattattattattatta
ttattattattattgctattgttattattattcttattattgctattgttattattattattcttattattgttgtt
gttgttgttcttattattgttgttgttgttattcttattattgtttattattattgtttttttttattctctaatta
ttccagtaatccataataaaaaataataaagtaaataaatagtaaatagtaaataattccagtaactgtagtaatac
acaataatctctaagaattaaaattgcattttgtaatgaaatatgttgattgttcgaatagttcagaaaaacttaaa
aatgcctcagcattaaacagttttgaggttgttcagggcatttagtttagatattttagtattttaaagcatttgtt
ttcattactacaaaaaagcaaatttatgagtgaattactttcagttcttctaaacgcctatgtgtatgcaattacat
aacaatagctctctttttttattgcatttttccttagtaatctaaatccaatctcttctttccctcttgcagATTAAA
GTCGGCAACGTGTACCCGATGACGTTGGAAATGTTTCAAAAATTGCTCAACGTGTCCTACTCCTATTTCACACTGCT
GCGCCGAGTGTACAACTAAACTTAACCGGTAAACAAACAAAAATCCCCTCATCACTATGCAAAGACAGCAAGCAGCC
GATCATCAAACACCATTAGCAGCCACAAAGTTACCAGCCGCTTATCCCACGGGATTTGGTGGAAAGTTATTGCACTG
AAGCTCTTTCACCCAAATTTTCATGGAGGTTCCCTCTCAACCAACCCATTGAAGCGAATAAAAGTATCAGCAACCAG
GCGACGGTGAAAAAACGCTGCATTATTGTGCTTGCTTCAGCATTCCAGCGAATGACTCTTAAACTTTTCCATTCAAA
AGTCGCGATGCTCACGATACGGAGCGGTGTGTTGTTCGATCCGCCGAGTGCACTCGCAAGCCGGTGATGTTGCCGGT
GGAAATGCACAGATCGACACAGCGATAGATAATCGTTTGTTCGCGTAAATGGGAGGGAAAAAAGTAAGCTGCCAGCT
ACTTCATTTCCATGTTAATTGAAACTCAAGCCAACGAACATGCAGAACCCGGTTGGTTGTGTGTCTCCGCTCCGGGA
AAGGTCTCTGCTCCGGGGCATGGATTCTTTCCCCCTCCGGGTGGTTGGGGGTATTGTTTAGGTTTTTATTTTACAAA
TTCATATCCTTCCGCTTCCGCATCAGCCGACCCGGTGGGTGCGCCAGACAGATGTGCGGCGGGCAACAAAACTATGC
ACGAACATGGCCAACAAACACAGCTTCTATCTCATCTCTGTGTCGCACTGTCTCGCTTTCCCGCTGCGTTGCTTGTA
GTACTATCATTGTTTTAGTCCACGGGTTTACTTCTAATTCCATTGCACCACGCAAAAAGGCTCATCCTTTGCTCGTT
CCGGTTGCAACTTCGACAAGCGCATGGTTGGGATACGAACAAAAAACCAACTACTCCACCCACTACTACTACTACTG
CCACCACCACTAACAACACTACACTTGGTTGGGAGCTTGCAGACCCACAAGCAAACAACGATACAAGCTAGCTAGCT
GCTGTGTGCGCTCGAGTCAGCCGACGGTACAAGGTTTAACCGGTACAAGCAACTCCCGGACCGATCCCAAAACTCTG
ACAAGGCACGGGGCCGCATCCGGCAGTACGGTCGGAAAACATGGAAATGTTTAATTAAAACTGTAATTGTCAATCGC
TGCTACAAGTTGTGACACAGGGAGAGAGAGAGACAGAGCGCGCCCGATGGTGATGGTGTAAAAGATAGATACAGGAA
AAGAGCGAGAAACATTGGTACGATTTGGTGTGGTTAGCAAATTTGATTTCCACTGATTTTGAGTGCAAATTTAATGC
ATCGAAAATTTGCCATTCAGGGTAAAGTTGCTCGTGGACGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT
CGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGGA

SEQ ID NO:11

Genomic Nucleic Acid Sequence
2083 nucleotides
Mosquito odorant receptor 3

AAGCAGAACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAATAGATGCC
TTCTGAGCGGCTTCGTCTCATTACTTCCTTCGGAACTCCTCAAGACAAACGCACGATGGTACTGCCAAAATTAAAGG
ATGAAACAGCAGTGATGCCGTTTCTGCTGCAAATTCAAACCATTGCCGGACTGTGGGGTGACCGTTCCCAGCGGTAC
CGTTTTTATCTCATCTTTTCCTACTTCTGCGCGATGGTGGTTCTACCCAAAGTGCTGTTCGGTTATCCAGATCTCGA
GGTTGCGGTACGCGGCACGGCCGAGCTGATGTTCGAATCGAACGCATTCTTCGGCATGCTAATGTTTTCCTTTCAAC
GCGACAACTACGAGCGATTGGTGCATCAGCTGCAGGATCTGGCAGCTCTAGgtgagtatgcagccaatcgattgttc
caaaccttcgcaacatccttcgtaacactgctacactttcagTCCTCCAAGACCTACCCACAGAGCTGGGAGAGTAC
CTGATCTCAGTGAACCGACGGGTCGATCGGTTCTCCAAAATTTACTGCTGCTGTCACTTTTCCATGGCAACGTTCTT
TTGGTTCATGCCCGTCTGGACGACCTATTCCGCCTACTTTGCTGTGCGCAACAGCACGGAACCGGTCGAGCACGTGT
TGCACCTCGAGGAAGAGCTGTACTTCCTGAACATTCGGACTTCGATGGCGCACTATACGTTTTATGTGGCCATTATG
TGGCCCACGATCTATACGCTCGGGTTTACCGGTGGCACAAAGCTGCTGACCATTTTCAGCAATGTTAAGTACTGTTC
GGCCATGCTGAAGCTCGTTGCACTCCGAATCCACTGTCTAGCGAGAGTAGCGCAAGACCGAGCGGAAAAGGAGCTGA
ACGAGATTATTTCCATGCATCAGCGGGTACTCAAgtaagtaaattcaaattgaaagttttgcagggaataacttgag
tgtgtctgacccgtgcacatcctagCTGCGTGTTCCTGCTGGAGACGACATTCCGCTGGGTATTTTTCGTGCAGTTC

ATTCAGTGTACAATGATCTGGTGCAGTCTCATCCTCTACATAGCGGTGACGgtaatagcattttcgtcatttcgtta
gccttattcaatccatttttgtgaacgtgaatttcccccagGGGTTCAGCTCGACGGTAGCGAATGTATGTGTCCAG
ATCATTTTGGTGACGGTGGAAACTTACGGCTACGGCTACTTCGGAACAGATCTAACCACGGAGGTGCTTTGGgtacc
ctttggatgaagcttcaaaaagtaattccaaattctgttttcgatttttccccttttccactagAGCTATGGCGTTG
CCCTCGCCATTTACGATAGCGAGTGGTACAAGTTTTCCATTTCGATGCGCCGCAAACTTCGACTGCTACTGCAACGA
TCCCAAAAACCGCTCGGCGTAACGGCGGGAAAGTTTCGCTTCGTCAATGTGGCCCAGTTTGGCAAGgtaacattaat
tacagtttgaaaattctgaagaatgcatcttacttgccttacttgttgttccagATGCTCAAGATGTCCTATTCATT
TTACGTAGTACTGAAGGAGCAGTTTTAGGAGCTGCTGTTTCCCACCCTGGAAATGGCCTTTTCGCACTGTCTTCTGT
TTGTTGGACGCACGCAGCACCGAGAGCGCCCCTGCACGCACTGACGTATTTTGGCTACTTTGACGTTTGCACCTTTG
ACAGCTGAAGGACAGGGTACAATTTTTGCTGCTGTTATTACGCGCAGCGCATTGGATACGAAAACATTGGCCACAAG
TTCTACGATTTTAGCGTTTATTTACTGTTCGTAGCAGCTTTTTTCCaCAATAAACACACACAATAACGTACCGACAG
TATTCTTTTCATTGTAGGATAGAGAAGCCGCCGGCCAGCAGCCAAAACGCGCCGCAAAACGAAAGGCGGCACCACCG
GGGGAAAAACACGGGAGCAAAACGAGAACAGAACGCAGTAAACAACAAAACCGGCCGGAACAACAACGGTGCCGGAA
ACGA

SEQ ID NO:12
Genomic Nucleic Acid Sequence
2374 nucleotides
Mosquito odorant receptor 4

GGGGAACTCCCCCACCCGACCAGACGACGGAAAGCTAACGATGTGCAATTGAA
TAGTCATTAGTAGCGTTTTTGCTCGCAAACGAACTAACCCTTTGACTTTTTAAG
TTCACTACGGTGAGGACAAAAATCAATAAATTAAATCGAGACCGTTGATGAGCA
AAAGAAAAAAAAAATATTTTACTGATTTTCATTTCGTTCCATCGACTACATAATCA
TAATTATATGCCACATTTTATTATAAGTTTTTGTATCATTTTTAAACAACACAAA
AATGCATCCTTTCGAATATTAGTCAGGTTGTATCAACA**ATG**AAGTTTGAACTGT
TTCAAAAATATTCCTCCCCGGACACGGTCTTATCCTTCGTGCTAAGGCTTTTGC
ATATCGTGGGCATGAATGGGGCAGGATTTCGGTCGCGAATTCGAGTTGGTGGC
ATTTTTCTGTTCTATTTAATCTTTCTTGTAATACCGCCACTAACGGGCGGGTAC
ACCGATGGTCACCAGCGTGTACGCACCAGTGTGGAATTCCTGTTTAATTGCAAT
ATTTACGGCGGCAGTATGTTCTTTGCCTACGATGTGGCCACTTTCCAAGCGTTC
ATCCAGGAACTGAAGAGCCTTTCGGTTTTGGgtaatatttaattaattaaaattgcgtttattgca
tcatcatttgtttctctttgcagTATGCTCACATTCGTACAGACTAAAGTATAAGCTGACCCG
GTTCAACCGTCGAGCGGATATTATCGCCAAAGTGCAAACGACCTGCATGGGTG
CTGTAACGCTTTTCTACTGGATTGCACCGATACCTTCCATCTGTGCGCACTACT
ACAGGTCGACCAATTCCACCGAACCCGTGCGGTTTGTGCAACATTTAGAGGTG
AAGTTCTATTGGCTCGAGAATCGCACCTCAGTCGAGGACTACATAACCTTCGTG
CTGATCATGCTACCCGTCGTGGTTATGTGTGGTTACGTATGCAATTTGAAGGTG
ATGACCATCTGCTGCAGCATTGGACACTGTACACTGTACACCAGGATGACTATA
GAGATGGTAGAGCAGTTGGAAAGCATGGCATCAGCGGAACGAACTGCCAGCGC
CATACGCAACGTGGGGCAGATGCACAGTGGTTTACTGAAATGCATTAGGCTTT
TGAACACGTCAATCCGATCGATGCTGATGCTGCAGTGGTTGACCTGCGTGTTA
AACTGGAGCATTTCTCTCATCTATCTAACGAACGTGgttagttttgtcttgtttggaaatccaa

aaacaaaaagatggctataattgaactttctattacagGGCATCTCGCTACAATCGGTTACCGTGG
TGGTAATGTTTTTTCTTGCCACTGCGGAAACTTTCCTGTATTGTTTACTTGGGA
CGCGGCTTGCGACACAACAGCAGCTGCTGGAGCACGCACTCTATGCTACACGG
TGGTACAACTACCCAATAGCCTTTCGCAGCAGCATTAGGATGATGTTGAGACA
GTCGCAAAGGCATGCACACATAACGGTGGGGAAGTTTTTTCGCGTTAATTTGG
AAGAATTTAGCAGGATTGTCAACTTATCCTACTCTGCTTACGTCGTACTTAAGG
ATGTAATAAAGATGGATGTACAG**TGA**ATGTTTTTTTTTTTTGGCTTGGCAACGAA
TGAAGTTTTCCGAATCTATATTAGATCTAGAATTTAATCTAGATGTCATAATATG
ATCTTGGCCATGACCGGTTCCTGGTTTTGGAACCAATTCTCAAAACAATTTTGA
ACTTAGGGCGAGGCATGAAATGTCCCAAGAACCTATCCAAGTTCTGGAACTAC
ATATTACCGAATCTATCCCATTATTGCCTCGGAACTGGTTTGGTGCTAAATATT
TGTCCAAATGTTGGTCCTGGACCTATCCAGACAAAGATCTTCAATTATTCCTAC
CACTGGAACTGATTAATTGATGTAGGAAGTCATGGAGGTGTTCAGGGAGAATT
TAAACACTAATGTTCCAACTCATTATTTCAAGGGCAATTCTATTTTTTATATGCC
CCTACGGATTGATACGTATGTATTACTCCATTTCCTGGACTTTGTCTTATTCTTG
CTGCTGATTGGACGTGAAATGTTGAGAAAAAGATTCTTATTTATGAGTGATACA
GAGCCTTTAAATACTCCTACGTTGTTTGCTATTTAAGTATGGCCAGGCTAATCA
CAATCGCTACTAATGAACAGAATCTCTTCTAATTAAACCCTTTCGATTGATAGT
GTCAATGTCAATGTCGAGATAATTGAACTGCAAACgATACCTACCTTAAACGGA
GCAGAACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGC
GAGGAGGAATAAAATAG

SEQ ID NO:13
cDNA Nucleic Acid Sequence
1194 nucleotides
Mosquito odorant receptor 4

ATGAAGTTTGAACTGTTTCAAAAATATTCCTCCCCGGACACGGTCTTATCCTTCGTGCTAAGGCTTTTGCATATCGT
GGGCATGAATGGGGCAGGATTTCGGTCGCGAATTCGAGTTGGTGGCATTTTTCTGTTCTATTTAATCTTTCTTGTAA
TACCGCCACTAACGGGCGGGTACACCGATGGTCACCAGCGTGTACGCACCAGTGTGGAATTCCTGTTTAATTGCAAT
ATTTACGGCGGCAGTATGTTCTTTGCCTACGATGTGGCCACTTTCCAAGCGTTCATCCAGGAACTGAAGAGCCTTTC
GGTTTTGGTATGCTCACATTCGTACAGACTAAAGTATAAGCTGACCCGGTTCAACCGTCGAGCGGATATTATCGCCA
AAGTGCAAACGACCTGCATGGGTGCTGTAACGCTTTTCTACTGGATTGCACCGATACCTTCCATCTGTGCGCACTAC
TACAGGTCGACCAATTCCACCGAACCCGTGCGGTTTGTGCAACATTTAGAGGTGAAGTTCTATTGGCTCGAGAATCG
CACCTCAGTCGAGGACTACATAACCTTCGTGCTGATCATGCTACCCGTCGTGGTTATGTGTGGTTACGTATGCAATT
TGAAGGTGATGACCATCTGCTGCAGCATTGGACACTGTACACTGTACACCAGGATGACTATAGAGATGGTAGAGCAG
TTGGAAAGCATGGCATCAGCGGAACGAACTGCCAGCGCCATACGCAACGTGGGGCAGATGCACAGTGGTTTACTGAA
ATGCATTAGGCTTTTGAACACGTCAATCCGATCGATGCTGATGCTGCAGTGGTTGACCTGCGTGTTAAACTGGAGCA
TTTCTCTCATCTATCTAACGAACGTGGGCATCTCGCTACAATCGGTTACCGTGGTGGTAATGTTTTTTCTTGCCACT
GCGGAAACTTTCCTGTATTGTTTACTTGGGACGCGGCTTGCGACACAACAGCAGCTGCTGGAGCACGCACTCTATGC
TACACGGTGGTACAACTACCCAATAGCCTTTCGCAGCAGCATTAGGATGATGTTGAGACAGTCGCAAAGGCATGCAC

ACATAACGGTGGGGAAGTTTTTTCGCGTTAATTTGGAAGAATTTAGCAGGATTGTCAACTTATCCTACTCTGCTTAC
GTCGTACTTAAGGATGTAATAAAGATGGATGTACAGTGA

SEQ ID NO:14
Amino Acid Sequence
412 residues
Mosquito odorant receptor 4

MKFELFQKYSSPDTVLSFVLRLLHIVGMNGAGFRSRIRVGGIFLFYLIFLVIPPLTGGYTDGHQRVRTSVEFL
FNCNIYGGSMFFAYDVATFQAFIQELKSLSVLVCSHSYRLKYKLTRFNRRADIIAKVQTTCMGAVTLFYWI
APIPSICAHYYRSTNSTEPVRFVQHLEVKFYWLENRTSVEDYITFVLIMLPVVVMCGYVCNLKVMTICCSIG
HCTLYTRMTIEMVEQLESMASAERTASAIRNVGQMHSGLLKCIRLLNTSIRSMLMLQWLTCVLNWSISLIY
LTNVGISLQSVTVVVMFFLATAETFLYCLLGTRLATQQQLLEHALYATRWYNYPIAFRSSIRMMLRQSQRH
AHITVGKFFRVNLEEFSRIVNLSYSAYVVLKDVIKMDVQNVSYSYFTLLRRVYN

SEQ ID NO:15
cDNA Nucleic Acid Sequence
1176 nucleotides
Mosquito odorant receptor 5

ATGGTGCTACCGAAGCTGTCCGAACCGTACGCCGTGATGCCGCTTCTACTACGCCTGCAGCG
TTTCGTTGGGCTGTGGGGTGAACGACGCTATCGCTACAAGTTCCGGTTGGCATTTTTAAGCTT
CTGTCTGCTAGTAGTTATTCCGAAGGTTGCCTTCGGCTATCCAGATTTAGAGACAATGGTTCG
CGGAACAGCTGAGCTGATTTTCGAATGGAACGTACTGTTTGGGATGTTGCTGTTTTCTCTCAA
GCTAGACGACTATGATGATCTGGTGTACCGGTACAAGGACATATCAAAGATTGCTTTCCGTA
AGGACGTTCCCTCGCAGATGGGCGACTATCTGGTACGCATCAATCATCGTATCGATCGGTTT
TCCAAGATCTACTGCTGCAGCCATCTGTGTTTGGCCATCTTCTACTGGGTGGCTCCTTCGTCC
AGCACCTACCTAGCGTACCTGGGGGCACGAAACAGATCCGTCCCGGTCGAACATGTGCTAC
ACCTGGAGGAGGAGCTGTACTGGTTTCACACCCGCGTCTCGCTGGTAGATTACTCCATATTC
ACCGCCATCATGCTGCCTACAATCTTTATGCTAGCGTACTTCGGTGGACTAAAGCTGCTAAC
CATCTTCAGCAACGTGAAGTACTGTTCGGCAATGCTCAGGCTTGTGGCGATGAGAATCCAGT
TCATGGACCGGCTGGACGAGCGCGAAGCGGAAAAGGAACTGATCGAAATCATCGTCATGCA
TCAGAAGGCGCTAAAATGTGTGGAGCTGTTGGAAATCATCTTTCGGTGGGTTTTTCTGGGAC
AGTTCATACAGTGCGTAATGATCTGGTGCAGCTTGGTTCTGTACGTCGCCGTTACGGGTCTCA
GCACAAAAGCGGCAAACGTGGGTGTACTGTTTATACTGCTAACAGTGGAAACCTACGGATTC
TGCTACTTTGGCAGTGATCTTACCTCGGAGGCAAGTTGTTATTCGCTGACACGTGCTGCGTAC
GGTAGCCTCTGGTATCGCCGTTCGGTTTCGATTCAACGGAAGCTTCGAATGGTACTGCAGCG
TGCCCAGAAACCGGTCGGCATCTCGGCTGGGAAGTTTTGCTTCGTCGACATTGAGCAGTTTG
GCAATATGGCAAAAACATCATACTCGTTCTACATCGTTCTGAAGGATCAATTTTAA

SEQ ID NO:16
Amino Acid Sequence
391 residues
Mosquito odorant receptor 5

MVLPKLSEPYAVMPLLLRLQRFVGLWGERRYRYKFRLAFLSFCLLVVIPKVAFGYPDLETMVRGTAELIFE
WNVLFGMLLFSLKLDDYDDLVYRYKDISKIAFRKDVPSQMGDYLVRINHRIDRFSKIYCCSHLCLAIFYWV
APSSSTYLAYLGARNRSVPVEHVLHLEEELYWFHTRVSLVDYSIFTAIMLPTIFMLAYFGGLKLLTIFSNVK
YCSAMLRLVAMRIQFMDRLDEREAEKELIEIIVMHQKALKCVELLEIIFRWVFLGQFIQCVMIWCSLVLYVA
VTGLSTKAANVGVLFILLTVETYGFCYFGSDLTSEASCYSLTRAAYGSLWYRRSVSIQRKLRMVLQRAQKP
VGISAGKFCFVDIEQFGNMAKTSYSFYIVLKDQF

SEQ ID NO:17
Partial cDNA Nucleic Acid Sequence
474 nucleotides
Mosquito odorant receptor 6

TTATGCTTACCGGATGTTGCGATCGCGCACGTGCTTTTCCGCATACGCCAGTGCACACTTGAT
GGCGGTGGTGATGACGTCTGCTGCGCACCGTTTTCTGCTCGTGAGTCAGACCTTTTCATTTCC
TGCAATATCCTGTTTCTTTCCCGACCCCACAGACGGTTAGACGGATATATGCTGGTAAAGTTT
GTCCTCTTCATGCTGTGCTTTCTGATCGAGCTGCTGATGCTGTGTGCGTACGGTGAGGATATT
GTGGAATCGCCTTGGGGTGATTGATGCCGCTTACGGTTGCGAATGGTACCGGGAAGGGTCGG
TGGCGTTCCATCGATCCGTGCTGCAAATTATACACCGCAGCCAGCAGTCCGTCATACTGACC
GCATGGAAAATTTGGCCCATCCAAATGAGTACTTTCAGTCAGATCCTGCAAGCTTCCTGGTC
CTACTTTACCCTCCTGAAGACCGTCTACGGGAATAA

SEQ ID NO:18
Partial Amino Acid Sequence
157 residues
Mosquito odorant receptor 6

LCLPDVAIAHVLFRIRQCTLDGGGDDVCCAPFSARESDLFISCNILFLSRPHRRLDGYMLVKFVLFMLCFLIE
LLMLCAYGEDIVESPWGDZCRLRLRMVPGRVGGVPSIRAANYTPQPAVRHTDRMENLAHPNEYFQSDPAS
FLVLLYPPEDRLRE

SEQ ID NO:19
cDNA Nucleic Acid Sequence

1206 nucleotides
Mosquito odorant receptor 7

ATGGTGCTGATCCAGTTCTTCGCCATCCTCGGCAACCTGGCGACGAACGCGGACGACGTGAA
CGAGCTGACCGCCAACACGATCACGACCCTGTTCTTCACGCACTCGGTCACCAAGTTCATCT
ACTTTGCGGTCAACTCGGAGAACTTCTACCGGACGCTCGCCATCTGGAACCAGACCAACACG
CACCCGCTGTTTGCCGAATCGGACGCCCGGTACCATTCGATTGCGCTCGCCAAGATGCGGAA
GCTGCTGGTGCTGGTGATGGCCACCACCGTCCTGTCGGTTGTCGCCTGGGTTACGATAACAT
TTTTCGGCGAGAGCGTCAAGACTGTGCTCGATAAGGCAACCAACGAGACGTACACGGTGGA
TATACCCCGGCTGCCCATCAAGTCCTGGTATCCGTGGAATGCAATGAGCGGACCGGCGTACA
TTTTCTCTTTCATCTACCAGGTACGTTGGCGGAATGGTATTATGCGATCGTTGATGGAGCTTT
CGGCCTCGCTGGACACCTACCGGCCCAACTCTTCGCAACTGTTCCGAGCAATTTCAGCCGGT
TCCAAATCGGAGCTGATCATCAACGAAGAAAAGGATCCGGACGTTAAGGACTTTGATCTGA
GCGGCATCTACAGCTCGAAGGCGGACTGGGGCGCCCAGTTCCGTGCGCCGTCGACGCTGCA
AACGTTCGACGAGAATGGCAGGAACGGAAATCCGAACGGGCTTACCCGGAAGCAGGAAAT
GATGGTGCGCAGCGCCATCAAGTACTGGGTCGAGCGGCACAAGCACGTTGTACGTCTCGTTT
CAGCAATCGGAGATACGTACGGTCCTGCCCTGCTGCTACACATGCTGACCTCCACCATCAAG
CTGACGCTGCTCGCCTACCAGGCAACGAAAATCGACGGTGTCAACGTGTACGGATTGACCGT
AATCGGATATTTGTGCTACGCGTTGGCTCAGGTTTTCCTGTTTTGCATCTTTGGCAATCGGCT
CATCGAGGAGAGCTCATCCGTGATGAAGGCGGCCTATTCCTGCCACTGGTACGACGGGTCCG
AGGAGGCAAAAACCTTCGTCCAGATCGTTTGTCAGCAGTGCCAGAAGGCGATGACTATTTCC
GGAGCCAAGTTTTTCACCGTTTCGCTCGATCTGTTTGCTTCGGTTCTTGGAGCCGTTGTCACC
TACTTCATGGTGCTGGTGCAGCTGAAGTAA

SEQ ID NO:20
Amino Acid Sequence
401 residues
Mosquito odorant receptor 7

MVLIQFFAILGNLATNADDVNELTANTITTLFFTHSVTKFIYFAVNSENFYRTLAIWNQTNTHPLFAESDAR
YHSIALAKMRKLLVLVMATTVLSVVAWVTITFFGESVKTVLDKATNETYTVDIPRLPIKSWYPWNAMSGP
AYIFSFIYQVRWRNGIMRSLMELSASLDTYRPNSSQLFRAISAGSKSELIINEEKDPDVKDFDLSGIYSSKAD
WGAQFRAPSTLQTFDENGRNGNPNGLTRKQEMMVRSAIKYWVERHKHVVRLVSAIGDTYGPALLLHMLT
STIKLTLLAYQATKIDGVNVYGLTVIGYLCYALAQVFLFCIFGNRLIEESSSVMKAAYSCHWYDGSEEAKTF
VQIVCQQCQKAMTISGAKFFTVSLDLFASVLGAVVTYFMVLVQLK

SEQ ID NO:21
Genomic Nucleic Acid Sequence
2272 nucleotides
Mosquito odorant receptor 5

tctagacttgaacccatgacgggcattttattgagtcgttcgagttgacgactgtaccacgggaccacccgtttatcactatcactattaattaattataatatgc
ttttgtagcgatcagcctaccgggttttgtttctctggatatcttaagttcccatttgattatcaagatagaacaacaacttgtaccttaaataatcattacgtaccc
ttaatcaacctgtgcatcaaggagttttcgcgaaagcaaaaatccgattgtctgatgttgtcttgattccatccgattcgttactggttctgcaaaatcgtccaat
aatacggcaatgtccttatcgatgcttgaatcaacatcacattgtttgcatttcgtttttgcgtgcaaatatgttatttgcaaagaaggcaaggtaatgtgcttaa
gagtaaatacaattcgctgtccatttttgtccaccagtgtgccagaacccgtgcctttagtccttcgaatacatccgaccagtcagcaagcaagtgcatcA
TGGTGCTACCGAAGCTGTCCGAACCGTACGCCGTGATGCCGCTTCTACTACGCCTGCAGCGT
TTCGTTGGGCTGTGGGGTGAACGACGCTATCGCTACAAGTTCCGGTTGGCATTTTTAAGCTTC
TGTCTGCTAGTAGTTATTCCGAAGGTTGCCTTCGGCTATCCAGATTTAGAGACAATGGTTCGC
GGAACAGCTGAGCTGATTTTCGAATGGAACGTACTGTTTGGGATGTTGCTGTTTTCTCTCAAG
CTAGACGACTATGATGATCTGGTGTACCGGTACAAGGACATATCAAAGATTGgtgcgtgataatgattg
ataaaaggaacctttgagcaactcctatccctttcaagCTTTCCGTAAGGACGTTCCCTCGCAGATGGGCGACTATCT
GGTACGCATCAATCATCGTATCGATCGGTTTTCCAAGATCTACTGCTGCAGCCATCTGTGTTT
GGCCATCTTCTACTGGGTGGCTCCTTCGTCCAGCACCTACCTAGCGTACCTGGGGGCACGAA
ACAGATCCGTCCCGGTCGAACATGTGCTACACCTGGAGGAGGAGCTGTACTGGTTTCACACC
CGCGTCTCGCTGGTAGATTACTCCATATTCACCGCCATCATGCTGCCTACAATCTTTATGCTA
GCGTACTTCGGTGGACTAAAGCTGCTAACCATCTTCAGCAACGTGAAGTACTGTTCGGCAAT
GCTCAGGCTTGTGGCGATGAGAATCCAGTTCATGGACCGGCTGGACGAGCGCGAAGCGGAA
AAGGAACTGATCGAAATCATCGTCATGCATCAGAAGGCGCTAAAgtaaggtctgccggtatgttgtggataga
atacatttctagctgctttcagATGTGTGGAGCTGTTGGAAATCATCTTTCGGTGGGTTTTTCTGGGACAGT
TCATACAGTGCGTAATGATCTGGTGCAGCTTGGTTCTGTACGTCGCCGTTACGgtaactaaaagcactg
tagtgatctgtctgccacaccattcactgctgtgtcttgttttgtcactcttcccagGGTCTCAGCACAAAAGCGGCAAACGTGGG
TGTACTGTTTATACTGCTAACAGTGGAAACCTACGGATTCTGCTACTTTGGCAGTGATCTTAC
CTCGGAGGCAAGTTGTTATTCGCTGAgtttcagttacttttccgttcccctctaaccgtaccacttgtaccatttgtttgagacagagc
ttgagcgtagCACGTGCTGCGTACGGTAGCCTCTGGTATCGCCGTTCGGTTTCGATTCAACGGAAG
CTTCGAATGGTACTGCAGCGTGCCCAGAAACCGGTCGGCATCTCGGCTGGGAAGTTTTGCTT
CGTCGACATTGAGCAGTTTGGCAATgtatggggagaccttccactgtggcaagaaagattttctttattaatgcatcttttaatttaca
gATGGCAAAAACATCATACTCGTTCTACATCGTTCTGAAGGATCAATTTTAAaggggaactcccccac
ccgaccagacgacggaaagctaacgatgtgcaattgaatagtcattagtagcgtttttgctcgcaaacgaactaacccttgacttttaagttcactacggt
gaggacaaaaatcaataaattaaatcgagaccgttgatgagcaaaagaaaaaaaaatattttactgattttcatttcgttccatcgactacataatcataattat
atgccacattttattataagtttttg

SEQ ID NO:22

Genomic Nucleic Acid Sequence
931 nucleotides
Mosquito odorant receptor 6

aacacccatcttatcggcaaaattagtatttaccgtttgaaagcggcttcccttcctggctgtttctcactctctctctctgtctctcttattgatgccgtatgcg
ccgcgtgctataggctagTTATGCTTACCGGATGTTGCGATCGCGCACGTGCTTTTCCGCATACGCCAG
TGCACACTTGATGGCGGTGGTGATGACGTCTGCTGCGCACCGTTTTCTGCTCGTGAGTCAGA
CCTTTTCATTTCCTGCAATATCCTGTTTCTTTCCCGACCCCACAGACGGTTAGACGGATATAT
GCTGGTAAAGTTTGTCCTCTTCATGCTGTGCTTTCTGATCGAGCTGCTGATGCTGTGTGCGTA
CGGTGAGGATATTGTGGAATCGgtaaggcaccaggcggtgatgagcgagtcgcgagtaattgaagcttttgcttttaaaacacatca
gagCCTTGGGGTGATTGATGCCGCTTACGGTTGCGAATGGTACCGGGAAGGGTCGGTGGCGTT
CCATCGATCCGTGCTGCAAATTATACACCGCAGCCAGCAGTCCGTCATACTGACCGCATGGA

AAATTTGGCCCATCCAAATGAGTACTTTCAGTCAGgtgagttgccaattgattgccgtttgcgttaatatttcagtaagagt
gcgctctttcccttagATCCTGCAAGCTTCCTGGTCCTACTTTACCCTCCTGAAGACCGTCTACGGGAA
TAAgtaagcgcgagagagagagagagagcagtatcgttcacccttggatgaatcaatagatttctaatcatgaaccattgaaaaatgaatcaacatttt
cgctagttgcacaatattgtaccattctatacagcttcaccacgaccaagcgtttgttgcatcaggaccaaacacgtttcgacaagccgcgtcacctgctgg
c

SEQ ID NO:23
Genomic Nucleic Acid Sequence
11,103 nucleotides
Mosquito odorant receptor 7

ccgcccgggcaggtgacttacgcggtctgacttgctggtgcgctgctttgtacggcaaacggctacacaagcgaatcgaattattttcctatcacgctgcg
cttaccagcgcctgctggtaggcaaagaatgtgcaaagtttcatttggcttggttcgtctgctttgctgtgaacgtgtgcacggttgcatcgctaaggtttcgg
tgtgagccgagaagttgcagatcgaaatctctttgtgtgtgtgtgtgtgtgtgcagtgggaagcattgtgtttagtgagaagtgaaaagaaaagtgctgaaa
aatgcaagtccagccgaccaagtacgtcggccttcgttgccgacctgatgccgaacattcgggttgatgcaggccagcggtcaactttctgttccggctac
gtcaccggcccgatactgatccgcaaggtgtactcctggtggacgctcgcccATGGTGCTGATCCAGTTCTTCGCCATCCTC
GGCAACCTGGCGACGAACGCGGACGACGTGAACGAGCTGACCGCCAACACGATCACGACCC
TGTTCTTCACGCACTCGGTCACCAAGTTCATCTACTTTGCGGTCAACTCGGAGAACTTCTACC
GGACGCTCGCCATCTGGAACCAGACCAACACGCACCCGCTGTTTGCCGAATCGGACGCCCG
GTACCATTCGATTGCGCTCGCCAAGATGCGGAAGCTGCTGGTGCTGGTGATGGCCACCACCG
TCCTGTCGGTTGTCGgtatgtgtgtatgtgtgtggccgtttgggaaagtgtctttgcggcagaaccccaatctactgttacgcttgactgggtttt
tgttttttctcggtggagggacgggataaaatatctgaaagaataattgagtcaacccacaggggatgcaagacatcgcaggcagagagtttgggtttg
atttatcaccgcacaccgaatatcttcacggttcataagcttcaccgcggtgaaaagggaactccccatttccctgttttcttttttttcttcctctcgataaattac
tcatcgcttttcgttttttttttttgttgttgcttctttcttctttcatccctactagCCTGGGTTACGATAACATTTTTCGGCGAGAGC
GTCAAGACTGTGCTCGATAAGGCAACCAACGAGACGTACACGGTGGATATACCCCGGCTGC
CCATCAAGTCCTGGTATCCGTGGAATGCAATGAGCGGACCGGCGTACATTTTCTCTTTCATCT
ACCAGGTACGTTGGCGGAATgtcctgcgcgtcacagttggcagtcagtgagcggcaacacggcgaaaaaatgggactaaaaccgg
tcttcacagagccaacacattcctacagcaattgcataccttcgggcggtcgggactgggcaatgcagctacaacatcctcgcctaaagttatgcaattcg
agcgacaaatgttgccgtgttagggctttttgtgataatagtcgtttttttgtcctctcgcttatcaaactctatcaacggaggaaatccattttcgctacaatgcc
tacagctcaagtttcaaggtcaatcgagcgggtggggatcaacttttttattcattttgctaacgccccatcaacaaattctatgttctcaatggcaaagattact
gcccgcaccaatcgcccaacgaaacggcaaaagaaaagcgacgattatgaagatgtccaaaccattgcccgcccgacgctttatctgatgatttgcggg
atggcttttacttgtctgctactttcaggcacaaaaggaaatgaaaccagcgcaggctcgtttgccggcttgcggaggttcttcaggcactgaggctgagta
cttaaatcgaacgatttttacgattctggatccagttttatgatgtggcctgcattacagtggcaattataccctgatgttcatttcattgcattttgtaagtttgtgct
ggtaacgcccgtaacgattaattcttttcaaagagattctttcaaagagattcaaaatgtgtataacaaatgctaacgaatggaccgtacttggagggttgcg
gaaagtaacgttttaaaatattcatcacaatcctctgcaaacttgtgcttaattaattggtgcacaataagtttaaactgtggcggcagatgtgtcgctgtccgc
ttccttccttcccagcaagctcgtgcgaaataatttattccatcattttaatacagccgtttgtgcattttaattagcaaagcaatataaaaagcagctaaccatc
cccattaaaacaaagtgcttccgggcccaattgttatggcggtggaaagtaatggttttaccagtggaagtgtcctttcccatcgtgggtacttcgcgatattc
ttgtcttatacaagtgcatacagaaaaaaaggacaaatcctccttgctatggtctaaggccagcttcggtaccgcttccgcttcgggatgtcataaagtttgat
gggtgtttttaacattacttccgctcttaaccacctaatggacttttcatgcttgagctaaagttaaaccagccaccagcggtacgcaccgagccacggttgat
ttcggcggcggcctcatccccagttttgcgccaccaatattgccttcattaatctgtaccctcggagcgttagggcccgcggacgagtcctcgttgtaatgc
accgccatgccacgggacgggataatccgttgggacggcgcgaaagcgactatcgcggacggattggttcgaccgtgctacaacacattttatgcttca
cagatttacttcctgctgttttcgatggtccagagcaacctcgcggatgtcatgttctgctcctggttgctgctagcctgcgagcagctgcaacacttgaaggt

aggtacggtagcaaacgtggttgtctttacatccgcgtgcagcattatccttatcgacgtgtagtgttaacggtaaaagaggaagcgataaaaaagcaaca
ttctctcacaccctcgatctctctttattttctctctctctctctctctctctctctctctctctctctctctctctctctctctccatctcctcgggcagGGTATTA
TGCGATCGTTGATGGAGCTTTCGGCCTCGCTGGACACCTACCGGCCCAACTCTTCGCAACTG
TTCCGAGCAATTTCAGCCGGTTCCAAATCGGAGCTGATCATCAACGAAGgtatgtgaaacgtgtgctcgtg
gcagacggactcaaagagagcataacacaatcccctggtagttcatttcaatgaccttaacactcggcaagctaagcgagacagtggggacagtgagaa
agagagaacaagaaaaaaaaccatcatccgtacgacatcatcgctacgtaccggtatttcaggatgaggaaataaaacgctaggggaatgaaagtgcg
acagaatgataaaacaatccccacccaggcccccagcctggacgaacggatgtagtgtgcgaagcgagcaaaaaaagtcaaataaattgaagtttaaa
aatagattttccccgtccatccgtggtggagcgtaaagcccggcggacaacttcgagcacggcgaccgtgcacagtactgtgccacagttgtagggacg
gataagctccgttccttttttatcctttttttttggagatttgtttgcgttcgcatcgttagacgagcttagtgccgtgttgctctaattgctatttattataaagcgctt
ccaaatagaagatcggttctctccatttaatctatcgcgcctgtacgcctgaaactatgcactgtgctgtgaaaccgtcaagctcgagcacgacgaatggcc
caccgtaccacgcccgtggtgcccaaagcgcaacgcgaattgcatgttaacaaacctttgcctaccatccaatccgtgtgaaattgcccgctctctttctct
cttttgcgctttcggtgtatcgaacggttttgtccctttttttttactttgctcttgatctcttgctgtgctcactttcatctcatgttttgcctgacggtggtgggttttcg
aaaaaagagcgatttcttctgcgtgtgtgtgtggttttttaaataaccgctccaggtcgtgttgaacgctgcaggaccgatcggagctagtttattatcagctt
tagtgtttatcccacccatgccccacatcacgtctgtggagagtgggggaagcttaagtccaatgtaatttaccgtgtttctgtcgttcgtcaccttcttcgtcg
atggagattggtgcggttggcacgataaaagcccactgcacgttacggaccgagggaaaggtctttttgtaggcctagcaacggtcctcattcaccgcat
gggggtgtagctcagatggtagagcgctcgcttagcatgtgagaggtaccgggatcgatacccggcatctccaacccacacaaaacgtttttaagaaga
ttttagggaagatattaacgcgggtacactgtgctcctctaagttggaagagtagatgagatgatgacaagggagaaggaacatgtgtacgtgtttgatag
caaacacacaaacaacaatatcatctctgataataatctgatgtgtgatgtgtgtgtattgttgttatgctgcctttgccatcttgtccctctctctcctgttcaact
cctaaaagaattgtttggagtcctctcagttcctcgtaaagatcctttcgagattcttctttccttttttattatttattccacgagcctctgacataagtagccttccg
cttatttccttctccttgcacttgtcagttccgtgtagagcgtcattttgaggtttacacatttcccaccgacgcctgattgttacattgtcatctacattgctttccgt
ttaccgttccgccctttttttttaacgctaccacagAAAAGGATCCGGACGTTAAGGACTTTGATCTGAGCGGCATC
TACAGCTCGAAGGCGGACTGGGGCGCCCAGTTCCGTGCGCCGTCGACGCTGCAAACGTTCG
ACGAGAATGGCAGGAACGGAAATCCGAACGGGCTTACCCGGAAGCAGGAAATGATGGTGC
GCAGCGCCATCAAGTACTGGGTCGAGCGGCACAAGCACGTTGTACGgtaggtatggtaatttctaaggtgtg
gtgtaaagcctccaggttccatgaaaaagggatactttaccacagtaagagtttgttttgctggacttacattctttggagcattgtttggtgttgtgctgaaacc
ggttgcaatatcgttttgcgaagaaattatgtgtaaagcgtattacaatctcattcctctgttaatctgtaccaattgtgtcagccccgaccgaaagcaggccta
attcgtaccagaaaaaccacaagctgtttgtaagcatcgatacgcccgaagctttcaatccagccaaggcgccacctactattgacgtgacttttgcacgtt
cacactctccctctcccattctttctataaccaatcgtcgctcagccagcatcgcccggagtgaagttttatttgaacgatatcacccgtatcgattttccacta
aacatgcttaaatcgtttcacaaagctcccccaaaatcccatttcaccaatccaccaatttgaagtccgtcgtccttgtgtccttgtgtttgtgtgtttgtgtgag
ctggagacatgggggagtgagtaaccgaacaacctcttgccgctgcttcacgatatcgaacagcaccaagataagcatccctttttccctagccgatgtct
ccgatatctcgattccgcttccagcgaggcaaagaaaaaggcgaactggctgacctcacccggggcgaggaaaaagcgtagggattacgtcgagcag
cacgagttgtgatttcttcttcttctggttccataaatcgctgacggtttccattaccgcctgcggagtgcacacacgtgaagggaaagcgaaaacgtttaga
ttccagcagcaacggcagcaccagaagcagcagcagcgcggcaaattgaatcatcctgacgcgatgagttgtctgggttttcgggtcggtggcttacag
caccacaccatctgctgcagctaatacagctgtaaatttcgttagacatagacttgattttacaatattacacacacacttacacacacagctatagatttgtcg
cttggcgtatggctctgtacggcgtgccgtacatgccgcgagccgtgttgctgctggttgcgatacggatcacgtccgattcgattcagcctgcgtgttttg
gtgaagatccttatcggtgacccactttcagtgtgtcgagagcgagggtcactatggcgcctgtcagttggaaagctaggctcgattcaaagggccattgt
gccagtgttcttttaagatagcgataagctttgatcgaaatagtaaatcaaacattgtttcttttttcctattccaaactgttgccaacctcattattacgttttgc
agcgggtgtatagtaaattgcatactttaaggcgtgattttcaaatgtagcgttccgtatgcagaaacgccatggattatgcaatttaaacaatgctgcttcctt
aacattcaaataacggcttattaaggaacttttgtgcaatttgtttttaacagcaaatagttagctcagaacgatcacatttagtatcgcttcaacaaagaactct
tttaaacacacaatttgtaatgccattccctcgagaaagtttcttgtcagtcctcctctgcatcacagcaacaaccaaacctgctcatgtttcctgctcgtttcct
agctgttttgaacgttatttccgattcctgtgcttgcccgcttttcttacaatcaaccacaatggttcagatttcgctcttatttttattgacccactgctttcgtgctga
agcccgtggaaacaatgcgccaagctcagcatccagccatgcatgtaaaatgagccacgcgacagattttagacatcgctttcgctctgcaccggaggt
ggttttattcttgtttccgattcccacgtccattcgtcctgggtccgtccgccgggcccgaaaccgtaagccgtgcggggaattacgcaatcgaaacgagc
cagaaaatgagcacgccaaatgcaaagaaaatccccctttgagtggtgctcctgccaccactcatctccccaactggtgggtgaaaaaccttgtgcgccc
cttctctttccagaaaaaaaacgcctcgctcgcacaaaaacatgctcgcccggtgaagctgcgtatgtcgcagaagctcaaaccaacgccgccagcaag
catcaacaatttctattcaaacacccaacgcagcgcccaaaccgggtgcactgtactcagtagcgaagatgctcagattgtcccgtgcgctgctttcgatg
cccgtttcggagcgggaagccatcgcttgccaacgttggcgatgtcttttagccgtggatttgaattttctgaatatcacaggcgggcgcggtttgcctgca

aggttgttgcttcccacacgagcattgctttccgtaccgcggtggggcgagttttcaacgcaaccttctacaagcaacgccacaacgcctgggagcgatat
ttaacagaaacaagaacatcccgaacttcagcacatgccgtgatttgcctgttggaaaagcttttgtgagcgtgtgagttgaacgagctctatttttcccagcg
atgggtggcatttgtgtggcatgctatcgtcagcttttcttgaatctttacctctccattcgcctccattagtacacgcgtatggaaaatgggtgcaacggatca
gaacggattttccgcgacagacttaataaagggaaagcaacgcgtttttgcatgtgtagtgtttatgagctttatgccgttactttgcaattaaaaatagcaaa
aaataacagttttttttttgtaagcggattacaaagaatgtatcagaatattacgtgaaacattcatttcatgctgttaacgctcaaatagaatagttttgtaacacg
gattgcataccttgccggtatcggttacattttcgcctaacagtatgcaatctgtttagctttgttgtttaatgactgcgttggtagtacaatatttatttacaccgc
gtaatttatctcacaaattgcaaaaaaatgtcaatctgtatcgattattcacacaaatcagatcccggaaccagtgtagcccaatgtgctcttattgaattacca
cgaacaaatcaacctgatgcccgggtccgttggcaaacagcttgcgccgaagccgctcagtgtttcgtgcactaccgtgctgccatttttgctgccctcatc
gaacagataaacagaagggcaactcttgtgagcatcgcaatgcccgtctgaagttccgtcgaaaatgggcctaaattcaatttgacgcatttacccgcgaa
caattgcgcgaaggctgtcaagtgtgttccacgaactgcgacaacaagcacacacacaaacacaaatgttatcgtttcggcatgtttctcggtacaaagcg
tgtggcgctatgtggcatgccgattcccagacagagtgatcgatagtaaatgtagcctatccggtagcattcaatttccttttctatcctcgcaaacaaagcc
cattctggggaggcgtggtgaagctttcaaaggcattgtgaaacaaatgtcctggttcggagggatgctggggaaagcaaacacggtgccgccatcgct
gctaccgtcaatcgatcatgcatgatgtgattaatatttgtgttattcacctgcgtatctatgcgtccgtcgtgtcgttcggatttccggaagtcaaggaaaaag
cgactccatttgggattggtttttgcagcgaaaaatcaaaacattcgcacaaaaccgtcctccatttcaaatgcctacacttgtcactgtatatctctctttctctc
gttttgccacgttgcagTCTCGTTTCAGCAATCGGAGATACGTACGGTCCTGCCCTGCTGCTACACATG
CTGACCTCCACCATCAAGCTGACGCTGCTCGCCTACCAGGCAACGAAAATCGACGGTGTCAA
CGTGTACGGATTGACCGTAATCGGATATTTGTGCTACGCGTTGGCTCAGGTTTTCCTGTTTTG
CATCTTTGGCAATCGGCTCATCGAGGAGgtacgtgcgctcggcgtgttgccgtgggaaagcattctccctgccccatatcgctt
cattctcccagatcacacatttgcatcacaaagccagcacacttttgcttcgccgctgccatctcggcttctgaatgttttcacttctcccatacttctcccgtgc
agAGCTCATCCGTGATGAAGGCGGCCTATTCCTGCCACTGGTACGACGGGTCCGAGGAGGCA
AAAACCTTCGTCCAGATCGTTTGTCAGCAGTGCCAGAAGGCGATGACTATTTCCGGAGCCAA
GTTTTTCACCGTTTCGCTCGATCTGTTTGCTTCGgtaagtgtagcctggtggctggcacagaacaggctggcaaaacagg
gactttggctctagcctgatgggtggtatatgtgtgtctatttttgctaccattctcgcatcccttcctttccagGTTCTTGGAGCCGTTGTCA
CCTACTTCATGGTGCTGGTGCAGCTGAAGTAAacagccgtggcccggaaggatgtgtttttttttcgctcgttcggttgtttgttt
gtgcacactttctcttggacattttctctactgcaaaggtttaacaaacagcaacaacaaataatcccaagttttcttttacagatctttgcaaaatgattagatttt
aatagattaacagtgcttgattatctgtcctgtagcaaccggggctgaagaacgttgatttggtaaaagtacaaaagggacgttggaaattgaaccaccag
aagagtgatatttatgcaaagctcaccaagggaaatctatgtatgtgtgatttgcgctcatcaagcactgtatgtgcctttcaactagtgcagcaataaagagt
acaaatgtttcttagcgcaccgtacattgtcgtttcggcgttttaaccgttgttgataatacacaaaagatgataaaaataaataataacaaaatgttaatatgag
taagtactaaatagagaaatcgttttagtatgatcatacctccaatcatttgtttgaaattaactttaattttaactcaaattaaaccgatgttttactttctgtgagaa
ttattgtggaagaacttaatggaagtataattaaattgattgctaactttatgcgtttttcaatttacgaacgctagtcttcaaacatcgcttcaaaagtattactac
cacattattcatttacttatagttatatttattgcctcttcatctttccatggccagaactactgcagaaaagcttctttttgctcgctttccgatggttggttggacg
aagttggtaacaaacggcaagcaattagcataaactattttcgcatcgagatggaaatgaatgtaccactagaaccgagtgaaatgaattacttttcaacttg
cacgccaaaaccattatctaaagtacgcacaacttaaaaacaaaccccaaattgtcgtccacccttcattccacttttcttgctacactttccgaccgagttctgt
agcgccagcagcaaaaaaatacatataaaaccttcatcactcaagctgtatcgagccagcgtgggttgtgtttgactgtgctgtgaaagaaagaagaaaa
aaaaaacacttccacgggaagctagcaattggaaatgcataaattaaccggaagaaattcgcaaaaccccgcaccgacgtaccgcaccgcatccgtac
cgataccggaacaaacggtgtgcgcgaaagaatccgctagcagccccactggcacgggtatttgcttttggttctgtgtttttcttccactggtttgggtgcc
tgggcgaaggctagctcggctactttcccggggccgcaattttctgcagcccaaggcggcgtgctcgtggggccaaaagaat